GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 02:50:15 ; Search time 1697.44 Seconds
(without alignments)
4205.899 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaaataccaata.....tcagggcaagtccccgtcaa 1395

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_ph:*
6: gb_pl1:*
7: gb_pl2:*
8: gb_pr1:*
9: gb_pr2:*
10: gb_pr3:*
11: gb_ro:*
12: gb_sy:*
13: gb_un:*
14: em_fun:*
15: em_hum1:*
16: em_hum2:*
17: em_in:*
18: em_om:*
19: em_or:*
20: em_ov:*
21: em_pat:*
22: em_ph:*
23: em_pl:*
24: em_ro:*
25: em_sts:*
26: em_sy:*
27: em_un:*
28: em_vi:*
29: gb_ba3:*
30: gb_in1:*
31: gb_in2:*
32: gb_in3:*
33: gb_pl3:*
34: gb_pr4:*
35: em_ba1:*
36: em_ba2:*
37: em_htg1:*
38: em_htg2:*
39: em_htg3:*
40: em_htg4:*
41: em_htg5:*
42: em_htg6:*
43: em_htg7:*
44: em_htg8:*
45: em_htg9:*
46: em_htg10:*
47: em_hum3:*
48: em_hum4:*
49: em_hum5:*
50: em_hum6:*
51: gb_pr5:*
52: gb_pr6:*
53: gb_pr7:*
54: gb_htg1:*
55: gb_htg2:*
56: gb_htg3:*
57: gb_htg4:*
58: gb_htg5:*
59: gb_htg6:*
60: gb_htg7:*
61: gb_htg8:*
62: gb_htg9:*
63: gb_htg10:*
64: gb_htg11:*
65: gb_htg12:*
66: gb_htg13:*
67: gb_htg14:*
68: gb_htg15:*
69: gb_htg16:*
70: gb_htg17:*
71: gb_htg18:*
72: gb_htg19:*
73: gb_htg20:*
74: gb_htg21:*
75: gb_htg22:*
76: gb_htg23:*
77: gb_sts1:*
78: gb_sts2:*
79: gb_vi1:*
80: gb_vi2:*
81: gb_pat1:*
82: gb_pat2:*
83: em_htg0:*
84: gb_htg24:*
85: gb_pr8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	1322.2	94.8	10531	1	AE002409	AE002409 Neisseria
	2	1320.6	94.7	1500	2	AX011576	AX011576 Sequence
c	3	1315.8	94.3	311321	2	NMA3Z2491	AL162754 Neisseria
	4	1263	90.5	1500	2	AX011574	AX011574 Sequence
	5	1258.2	90.2	1500	2	AX011572	AX011572 Sequence
	6	1061	76.1	1110	2	AX011578	AX011578 Sequence
	7	365.6	26.2	1961	2	PAU32853	U32853 Pseudomonas
	8	365.6	26.2	4587	2	PAU49151	U49151 Pseudomonas
	9	365.6	26.2	10977	1	AE004511	AE004511 Pseudomon
	10	274.8	19.7	2027	2	AVU30799	U30799 Azotobacter
	11	260.4	18.7	12245	1	AE004037	AE004037 Xylella f
	12	260.2	18.7	390	2	AX011580	AX011580 Sequence
	13	246	17.6	1362	2	AF293977	AF293977 Aeromonas
	14	217.4	15.6	11461	1	AE004859	AE004859 Pseudomon
	15	212.4	15.2	1974	2	RMU31512	U31512 Rhizobium m
	16	211.8	15.2	2009	2	BAU07352	U07352 Brucella ab
	17	197	14.1	1905	2	BRUHTRAH	L09274 Brucella ab
c	18	196.6	14.1	10566	1	AE004142	AE004142 Vibrio ch
	19	196	14.1	2714	2	ECU32495	U32495 Escherichia
	20	196	14.1	3156	2	ECU15661	U15661 Escherichia
	21	196	14.1	10713	1	AE000402	AE000402 Escherich

	22	196	14.1	110000	2	ECOUW67_1 Continuation (2 of
	23	186.6	13.4	1191	2	PAU29172 U29172 Pseudomonas
c	24	183.4	13.1	18424	1	AF163841 AF163841 Myxococcu
	25	181.2	13.0	1476	1	AB040030 AB040030 Shigella
	26	178	12.8	10199	1	AE000125 AE000125 Escherich
	27	178	12.8	82727	2	ECO82K D26562 Escherichia
	28	178	12.8	123171	2	ECU70214 U70214 Escherichia
	29	176.4	12.6	1854	81	AR095624 AR095624 Sequence
	30	176.4	12.6	1855	2	ECHTRA X12457 Escherichia
	31	176.4	12.6	1855	2	ECOHTRAA M36536 E.coli htrA
	32	175.6	12.6	1980	29	STSPHSPG X54548 S.typhimuri
	33	175.6	12.6	1980	81	AR037634 AR037634 Sequence
	34	175.6	12.6	1980	81	AR084295 AR084295 Sequence
	35	175.2	12.6	1980	81	A18802 A18802 htrA gene.
	36	164.6	11.8	4151	2	RCFDXE Y11304 R.capsulatu
	37	164.4	11.8	1800	2	D78376 D78376 Yersinia en
	38	157.2	11.3	1613	2	BAU07351 U07351 Brucella ab
	39	154.2	11.1	1791	81	AR000272 AR000272 Sequence
	40	154.2	11.1	1791	81	I51141 I51141 Sequence 7
	41	154.2	11.1	1791	82	I78681 I78681 Sequence 7
	42	154.2	11.1	1889	2	ROCHTRA L20127 Rochalimaea
	43	149.6	10.7	11201	1	AE001732 AE001732 Thermotog
	44	148.2	10.6	11881	1	AE001355 AE001355 Chlamydia
	45	147.6	10.6	2778	2	BJY13616 Y13616 Bradyrhizob

ALIGNMENTS

RESULT 1
AE002409/c
LOCUS AE002409 10531 bp DNA BCT 25-MAY-2000
DEFINITION Neisseria meningitidis serogroup B strain MC58 section 51 of 206 of the complete genome.
ACCESSION AE002409 AE002098
VERSION AE002409.1 GI:7225757
KEYWORDS .
SOURCE Neisseria meningitidis MC58.
ORGANISM Neisseria meningitidis MC58
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 10531)
AUTHORS Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V., Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.
TITLE Complete genome sequence of Neisseria meningitidis serogroup B strain MC58
JOURNAL Science 287 (5459), 1809-1815 (2000)
MEDLINE 20175555
PUBMED 10710307
REFERENCE 2 (bases 1 to 10531)
AUTHORS Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V., Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappuoli,R. and Venter,J.C.
TITLE Direct Submission
JOURNAL Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES Location/Qualifiers
source 1. .10531
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
gene complement(132. .1631)
/gene="NMB0532"
/note="This region contains an authentic frame shift and is not the result of a sequencing artifact; similar to PID:1184684 PID:1220194 percent identity: 66.59; identified by sequence similarity; putative;protease DO, authentic frameshift"
gene complement(1769. .2398)
/gene="NMB0533"
CDS complement(1769. .2398)
/gene="NMB0533"
/note="similar to GB:J02857 SP:P20625 PID:146972 GB:U00096 PID:1742691 percent identity: 86.12; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="endonuclease III"
/protein_id="AAF40962.1"
/db_xref="GI:7225758"
/translation="MNRHIRQEIFERFRAANPHPTTELNFNSPFELLIAVLLSAQATD VGVNKATAKLFPVADTPQAMLDLGLDGVMEYTKTIGLYKTKSKHIMQTCRILLEKYNG EVPEDREALESLPGVGRKTANVVLNTAFGHPVMAVDTHIFRVSNRTKIAPGKDVREVE DKLMRFIPKEFLMDAHHWLILHGRYTCKALKPQCQTCIINDLCEYPAKA"
gene complement(2444. .2857)
/gene="NMB0534"
CDS complement(2444. .2857)
/gene="NMB0534"
/note="conserved hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40963.1"
/db_xref="GI:7225759"
/translation="MIPSYTYAALAFSAFTSATLLPGTSEAAFALFVHRFPEHAYGAL LCAGLANGLGSMVSYWMGRLLPSRKMPSEKTLNLMRRFGIWLLAFTWLPVVGDALPLA AGWLRLNPWTSGLMLVIGKTARYAFILWGMQYYAA"
gene 3390. .4670
/gene="NMB0535"
CDS 3390. .4670
/gene="NMB0535"
/note="similar to PID:1171339 percent identity: 71.89; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="glucose/galactose transporter"
/protein_id="AAF40964.1"
/db_xref="GI:7225760"
/translation="MSISIFLKRVIFRHHERQIMSAQSQNNHTSPLVVLTTLFFMMGF ITCMNDILIPHLKEIFDLSYVQAMLIQFCFFTAYAVMSIPMGAFVGKVGYKNGVIGGF LLTAVGCLLFYPAAGSHSYAVFLGALFILASGVTLLQVAGNPYVTLLAKPGKESATLT LVQAFNALGTTIAPQIGAFLILADATQTVSKAEQISSVQIPYLGLAGLLIILAVFVKM IRLPDARKIAAEESAHNHDGKTSVWQYKHLVFGTAGIFCYVGAEVSIGSLMVNVLGYL KGLDHASAAHYLSFYWGGAMVGRFLGSAVMAKFAPNRYLAFNASAAVVLLAVAMATGS GNADVAMWSLLAIGFFNSIMFPTIFSLATKGLGKFTNAASGVLCTAIVGGAVVPVVQG WVADTYTLMSSFVVSVICYLYIVFFAVYGYRADK"
gene 4961. .6340
/gene="NMB0536"
CDS 4961. .6340
/gene="NMB0536"
/note="similar to GB:M73530 SP:P27611 PID:143245 PID:2209268 percent identity: 65.57; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="Na+/H+ antiporter"
/protein_id="AAF40965.1"
/db_xref="GI:7225761"
/translation="MFAFKSLLDMPRGEALAVVVALIAAMGYTIISLEWLPHMSIIAA IVVLILYGLARGLKYNDMQQGMIGALNQGMGAIYLFFFIGLMVSALMMSGAIPTLMYY GFGLISPTYFYFSSFALCSVIGVSIGSSLTTCATVGVAFMGMAAAFQADMAMTAGAIV

SGAFFGDKMSPLSDTTGISASIVGIDLFEHIKNMMYTTIPAWLISAALMLWLLPNVAA
QDLNSVESFRSQLEATGLVHGYSLIPFALLVILALMRINAVVAMLFTVMVAVAVTYLH
STPDLRQLGAWFYGGYKLEGEAFKDVVKLISRGGLESMFFTQTIVILGMSLGGLLFAL
GVIPSLLEAIRTFLTNAGRATFSVAMTSVGVNFLIGEQYLSILLSGETFKPVYDKLGL
HSRNLSRTLEDAGTVINPLVPWSVCGVFISHALGVPVWEYLPYAFFCYLSLALTLLFG
WTGLTLSKK"

gene complement(6469. .7293)
/gene="NMB0537"

CDS complement(6469. .7293)
/gene="NMB0537"
/note="conserved hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40966.1"
/db_xref="GI:7225762"
/translation="MDGAQPKTNFFERLIARLAREPDSAEDVLNLLRQAHEQEVFDAD
TLLRLEKVLDFSDLEVRDAMITRSRMNVLKENDSIERITAYVIDTAHSRFPVIGEDKD
EVLGILHAKDLLKYMFNPEQFHLKSILRPAVFVPEGKSLTALLKEFREQRNHMAIVID
EYGGTSGLVTFEDIIEQIVGEIEDEFDEDDSADNIHAVSSERWRIHAATEIEDINTFF
GTEYSSEEADTIGGLVIQELGHLPVRGEKVLIGGLQFTVARADNRRLHTLMATRVK"

gene complement(7301. .7810)
/gene="NMB0538"

CDS complement(7301. .7810)
/gene="NMB0538"
/note="conserved hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40967.1"
/db_xref="GI:7225763"
/translation="MKRTKKYPFLTLQRQRFHLNFENASSAAGIPAERDFYRWAWSAL
KNEYRRADISLILLDEEEARAYNRDYRGKDYATNVLSFALNEGEILPCQVSEKLYGDL
IICPQVVLKEAAEQGKTPEQHFAHLTIHGTLHLMGYDHIEDDEAEIMEAEEIRLMRAA
GFPNPYQED"

gene 7929. .8864
/gene="NMB0539"

CDS 7929. .8864
/gene="NMB0539"
/note="similar to PID:151255 PID:496214 SP:Q60169 percent identity: 72.96; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="porphobilinogen deaminase"
/protein_id="AAF40968.1"
/db_xref="GI:7225764"
/translation="MNPKKLVIASRESLLAMWQAKHIQGRLKALYPDCEVEILGMTTR
GDQILDKTLSKVGGKGLFVKELEQALYDGRADLAVHSIKDVPMDLPEGFALAAIGERA
NPFDAFVSNQYTRLEEMPEGAVVGTSSLRREAQLRARYPHLLIKPLRGNVQTRLSKLD
NGEYDAIILAAAGLQRLKLDGRIRMILSESDSLPAAGQGALGIEIAAHREDLYEVLKP
LNHGVTNACVTAERALARALGGSCQVPLAAYCTEENGLLTLRGLVGHPDGSVVLRADA
QAPAEYADALGRAVAKKLADDGARELIGAVLNTEN"

gene complement(9268. .10461)
/gene="NMB0540"

CDS complement(9268. .10461)
/gene="NMB0540"
/note="similar to GB:L42023 SP:P44425 PID:1007884 PID:1221762 PID:1205848 percent identity: 84.60; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="aspartate aminotransferase"
/protein_id="AAF40969.1"
/db_xref="GI:7225765"
/translation="MFFKHIEAAPADPILGLGEAFKAETRPEKVNLGIGVYKDASGAT
PLVKAVKEAEKRLLESETTKNYLTIDGVADYNAQTQILLFGKDHEIIASRRAKTAQSL
GGTGALRIAAEFAKRQLNAQTIWISNPTWPNHNAIAKAVGIQDKPYRYYDAAKHGLDW
DGMIEDLSQAQKGDIVLLHGCCHNPTGIDPTPEQWETLAKLSAEKGWLPLFDFAYQGF
GNGLEEDAYGLRVFLKHNTELLIASSYSKNFGMYNERVGAFTLVAEDEETAARAHSQV
KTIIRTLYSNPASHGANTIALVLKNDDLKAQWIAELDEMRGRIKAMRQKFVGLLKAKG

ASQNFDFIIKQNGMFSFSGLTPEQVDRLKNEFAIYAVRSGRINVAGITDNNIDYLCES
IVKVL"

BASE COUNT 2151 a 2498 c 2982 g 2900 t
ORIGIN

Query Match 94.8%; Score 1322.2; DB 1; Length 10531;
Best Local Similarity 97.2%; Pred. No. 1.2e-234;
Matches 1356; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

```
Qy      1 gtgttcaaaaaatacccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          |||||||||||||||||||| |||||||||| |||||||||||||||| ||||||||||
Db   1631 GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCCTCGCTGGCAGGC 1572

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1571 TGCGACAAGGCAGGCAGCTTTTTCGGTGCGGACAAAAAAGAAGCATCCTTCGTAGAACGC 1512

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1511 ATCGAACACACCAAAGACGACGGCAGCGTCAGTATGCTGCTGCCCGACTTTGCCCAACTG 1452

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
          |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1451 GTTCAAAGTGAAGGTCCGGCAGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGCACCCAA 1392

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          |||||||||||||||||||| |||||||||||| |||||||| ||||||||||||||||
Db   1391 AACGGCAGCGGCAATGCCGAAAACGATTCCGACCCGATTGCCGACAACGACCCGTTCTAC 1332

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatccccccaagaagaagcagatgac 360
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   1331 GAATTTTTCAAACGCCTCGTCCCGAATATGCCCGAAATCCCCCAAGAAGAAGCAGATGAC 1272

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1271 GGCGGATTGAACTTCGGTTCGGGCTTCATCATCAGCAAAGACGGCTACATCCTGACCAAT 1212

Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          |||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db   1211 ACCCACGTCGTTACCGGCATGGGCAGTATCAAAGTCCTGCTCAACGACAAGCGCGAATAT 1152

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1151 ACCGCCAAACTCATCGGTTCGGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCA 1092

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          |||||||||||| |||||||||||||||||||||||||| |||||||||||||||||||
Db   1091 ACGGAAGAGCTGCCCGTCGTCAAAATCGGCAATCCCAAAGATTTGAAACCGGGCGAATGG 1032

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   1031 GTCGCCGCCATCGGCGCGCCCTTCGGCTTCGACAACAGCGTGACCGCCGGCATCGTGTCC 972

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    971 GCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCC 912

Qy    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    911 ATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACTTAAAAGGACAGGTCGTCGGCATC 852

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
          || |||||||||||||||||||||||||||||||||||| ||||| ||||||||||||||
Db    851 AACTCGCAAATATACAGCCGCAGCGGCGGATTCATGGGCATTTCCTTCGCCATCCCGATT 792

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          |||||||,|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    791 GACGTTGCCATGAATGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAA 732

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          |||||||||||||||| |||||||||||||||||||||||||| ||||||||| |||| |||
Db    731 CTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTCGGTTTGGACAAA 672
```

```
Qy     961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
           ||| |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db     671 GCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCCCCGCAGAACGTGCCGGCCTG 612

Qy    1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttt 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     611 CAGGCGGGCGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCCGGCGACCTT 552

Qy    1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 CCCGTTATGGTCGGCGCCATTACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGGCGCAAA 492

Qy    1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
           |||||||||||||||||||| |||||||||||||||||||||||| ||||||||||
Db     491 GGCGAAGAAATCACAATCAAAGTCAAGCTGGGCAACGCCGCCGAGCATATCGGCGCATCA 432

Qy    1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     431 TCCAAAACAGATGAAGCCCCCTACACCGAACAGCAATCCGGTACGTTCTCGGTCGAATCC 372

Qy    1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
           |||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db     371 GCAGGCATTACCCTTCAGACACATACCGACAGCAGCGGCGGACACCTCGTCGTCGTACGG 312

Qy    1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
           |||||||||||||||||||||||||||||| |||| |||||||||||| || |||||| ||
Db     311 GTTTCCGACGCGGCAGAACGCGCAGGCTTGAGGCGCGGCGACGAAATTCTTGCCGTC-GG 253

Qy    1381 gcaagtccccgtcaa 1395
           ||||||||||||||||
Db     252 GCAAGTCCCCGTCAA 238


RESULT   2
AX011576
LOCUS       AX011576     1500 bp    DNA             BCT       18-SEP-2000
DEFINITION  Sequence 5 from Patent WO9955872.
ACCESSION   AX011576
VERSION     AX011576.1  GI:9998107
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1500)
  AUTHORS   Ruelle,J.L.
  TITLE     Basb013 dna and proteins from neisseria meningitidis
  JOURNAL   Patent: WO 9955872-A 04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
     source          1. .1500
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      385 a    471 c    382 g    262 t
ORIGIN


  Query Match             94.7%;  Score 1320.6;  DB 2;  Length 1500;
  Best Local Similarity   97.1%;  Pred. No. 3.4e-234;
  Matches 1355;  Conservative    0;  Mismatches   39;  Indels    1;  Gaps    1;

Qy       1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
           |||||||||||||||||| |||||||||| ||||||||||||||| ||||||||||
Db       1 GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCCTCGCTGGCAGGC 60

Qy      61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGCGACAAGGCAGGCAGCTTTTTCGGTGCGGACAAAAAAGAAGCATCCTTCGTAGAACGC 120

Qy     121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     121 ATCGAACACACCAAAGACGACGGCAGCGTCAGTATGCTGCTGCCCGACTTTGCCCAACTG 180

Qy     181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagccccgccccgcgcacccaa 240
           |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTTCAAAGTGAAGGTCCGGCAGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGCACCCAA 240

Qy     241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
           |||||||||||||||||||| |||||||||||| ||||||||| ||||||||||||||
Db     241 AACGGCAGCGGCAATGCCGAAAACGATTCCGACCCGATTGCCGACAACGACCCGTTCTAC 300

Qy     301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     301 GAATTTTTCAAACGCCTCGTCCCGAATATGCCCGAAATCCCCCAAGAAGAAGCAGATGAC 360

Qy     361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     361 GGCGGATTGAACTTCGGTTCGGGCTTCATCATCAGCAAAGACGGCTACATCCTGACCAAT 420

Qy     421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
           |||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db     421 ACCCACGTCGTTACCGGCATGGGCAGTATCAAAGTCCTGCTCAACGACAAGCGCGAATAT 480

Qy     481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ACCGCCAAACTCATCGGTTCGGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCA 540

Qy     541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
           |||||||||| |||||||||||||||||||||||||||| |||||||||||||||||||
Db     541 ACGGAAGAGCTGCCCGTCGTCAAAATCGGCAATCCCAAAGATTTGAAACCGGGCGAATGG 600

Qy     601 gtcgctgccatcggcgcgccccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
           ||||| |||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     601 GTCGCCGCCATCGGCGCGCCCCTTCGGCTTCGACAACAGCGTGACCGCCGGCATCGTGTCC 660

Qy     661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCC 720

Qy     721 atcaatccgggcaatttccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACTTAAAAGGACAGGTCGTCGGCATC 780

Qy     781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
           || |||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db     781 AACTCGCAAATATACAGCCGCAGCGGCGGATTCATGGGCATTTCCTTCGCCATCCCGATT 840

Qy     841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgccggacaa 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GACGTTGCCATGAATGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCCGGACAA 900

Qy     901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
           |||||||||||||||| |||||||||||||||||||||||| |||||||  |||| |||
Db     901 CTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTCGGCTTGGACAAA 960

Qy     961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
           ||| |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db     961 GCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCCCCGCAGAACGTGCCGGCCTG 1020

Qy    1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttt 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CAGGCGGGCGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCCGGCGACCTT 1080

Qy    1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CCCGTTATGGTCGGCGCCATTACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGGCGCAAA 1140

Qy    1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
           |||||||||||||||||||| |||||||||||||||||||||||| ||||||||||
Db    1141 GGCGAAGAAATCACAATCAAAGTCAAGCTGGGCAACGCCGCCGAGCATATCGGCGCATCA 1200

Qy    1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TCCAAAACAGATGAAGCCCCCTACACCGAACAGCAATCCGGTACGTTCTCGGTCGAATCC 1260
```

```
Qy   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
          ||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db   1261 GCAGGCATTACCCTTCAGACACATACCGACAGCAGCGGCGGACACCTCGTCGTCGTACGG 1320

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          |||||||||||||||||||||||||||| |||| |||||||||||| || |||||| ||
Db   1321 GTTTCCGACGCGGCAGAACGCGCAGGCTTGAGGCGCGGCGACGAAATTCTTGCCGTC-GG 1379

Qy   1381 gcaagtccccgtcaa 1395
          |||||||||||||||
Db   1380 GCAAGTCCCCGTCAA 1394
```

RESULT 3
NMA3Z2491/c
LOCUS NMA3Z2491 311321 bp DNA BCT 30-MAR-2000
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome; segment 3/7.
ACCESSION AL162754 AL157959
VERSION AL162754.2 GI:7379424
KEYWORDS .
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 311321)
AUTHORS Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N., Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
TITLE Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
REFERENCE 2 (bases 1 to 311321)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT Notes:
Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).

FEATURES Location/Qualifiers

```
     source          1. .311321
                     /organism="Neisseria meningitidis"
                     /strain="Z2491"
                     /db_xref="taxon:487"
                     /note="serogroup: A"
     RBS             115. .120
     gene            129. .587
                     /gene="NMA0698"
     CDS             129. .587
                     /gene="NMA0698"
                     /note="NMA0698, possible periplasmic protein, len: 152 aa;
                     unknown, contains a probable N-terminal signal sequence
                     and lies within a region of unusually low GC content"
                     /codon_start=1
                     /transl_table=11
                     /product="putative periplasmic protein"
                     /protein_id="CAB83985.1"
                     /db_xref="GI:7379425"
                     /translation="MLTKLKILLFLFLFVFVLAINLLFFFFSSDIESFGNYQFEYVYD
                     KGWPANYILVMKDGNEGNFDKIISGLVLEYYKEDDNIYFSYIDGQGFASDSCYYKPEI
                     LYGKIILNKNHIININSMEKNNFLSEDKIMKGTRNWLADPKNKCNIQTLD"
     gene            742. .1077
                     /gene="NMA0699"
     CDS             742. .1077
                     /gene="NMA0699"
                     /note="NMA0699, pseudogene, probable ABC transporter
                     protein, len: 336 bp; similar to C-termini of many ABC
                     transporters e.g. TR:Q46973 (EMBL:U47048), mtfB,
                     Escherichia coli microcin transport protein (707 aa),
                     fasta scores; E(): 2.6e-19, 53.2% identity in 111 aa
                     overlap. Contains PS00211 ABC transporters family
                     signature. NMA0686 may be the remainder of this
                     pseudogene"
                     /codon_start=1
                     /pseudo
                     /transl_table=11
                     /product="putative ABC transporter protein (pseudogene)"
     misc_feature    868. .912
                     /gene="NMA0699"
                     /note="PS00211 ABC transporters family signature"
     gene            complement(1101. .2327)
                     /gene="NMA0700"
     CDS             complement(1101. .2327)
                     /gene="NMA0700"
                     /note="NMA0700, possible ribonuclease BN, len: 408 aa;
                     similar to SW:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
                     coli ribonuclease BN (EC 3.1.-.-), fasta scores; E():
                     2.3e-28, 34.3% identity in 268 aa overlap. Longer than rbn
                     at the C-terminus. Also similar to TR:O85449
                     (EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
                     aa), fasta scores; E(): 0, 45.0% identity in 333 aa
                     overlap. Contains a region similar to NMA0677, fasta
                     scores; E(): 4.3e-08, 64.4% identity in 45 aa overlap.
                     Contains hydrophobic, possible membrane-spanning regions,
                     as does rbn"
                     /codon_start=1
                     /transl_table=11
                     /product="putative ribonuclease BN"
                     /protein_id="CAB83987.1"
                     /db_xref="GI:7379426"
                     /translation="MTFLQRLQGLADNKICAFAWFVVRRFDEERVPQAAASMTFTTLL
                     ALVPVLTVMVAVASIFPVFDRWSDSFVSFVNQTIVPQGADMVFDYINAFREQANRLTA
                     IGSVMLVVTSLMLIRTIDNTFNRIWRVNSQRPWMMQFLVYWALLTFGPLSLGVGISFM
                     VGSVQDAALASGAPQWSGALRTAATLTFMTLLLWGLYRFVPNRFVPARQAFVGALATA
                     FCLETARSLFTWYMGNFDGYRSIYGAFAAVPFFLLWLNLLWTLVLGGAVLTSSLSYWQ
                     GEAFRRGFDSRGRFDDVLKILLLLDAAQKEGKALPVQEFRRHINMGYDELGELLEKLA
                     RHGYIYSGRQGWVLKTGADSIELNELFKLFVYRPLPVERDHVNQAVDAVMMPCLQTLN
                     MTLAEFDAQAKKQQQS"
     misc_feature    1376. .1385
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     misc_feature    complement(2417. .2426)
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     misc_feature    2444. .2453
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     gene            2486. .2743
                     /gene="NMA0701"
     CDS             2486. .2743
                     /gene="NMA0701"
                     /note="NMA0701, possible pseudogene, len: 258 bp; shows
                     weak similarity to part of SW:WRBA_ECOLI (EMBL:M99166),
                     wrbA, Escherichia coli Trp repressor binding protein (197
                     aa), fasta scores; E(): 0.064, 29.3% identity in 82 aa
                     overlap. Also similar to part of TR:O85450 (EMBL:AF067083)
                     Vitreoscilla sp. Trp repressor binding protein (fragment)
                     (124 aa), fasta scores; E(): 6.6e-14, 56.6% identity in 83
                     aa overlap"
                     /codon_start=1
                     /pseudo
                     /transl_table=11
                     /product="hypothetical protein NMA0701 (pseudogene)"
     misc_feature    complement(2755. .2764)
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     misc_feature    2778. .2787
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
```

```
RBS             2861. .2864
gene            2871. .3530
                /gene="NMA0702"
CDS             2871. .3530
                /gene="NMA0702"
                /note="NMA0702, len: 219 aa; similar to TR:Q52925
                (EMBL:Z50189), exsB, Rhizobium meliloti putative regulator
                of succinoglycan biosynthesis (not a transcriptional
                regulator) (234 aa), fasta scores; E(): 3.1e-15, 32.7%
                identity in 205 aa overlap. Also similar to many
                hbacterial ypothetical proteins e.g. SW:YBAX_HAEIN
                (EMBL:U32798), HI1191, Haemophilus influenzae hypothetical
                protein (196 aa), fasta scores; E(): 0, 79.2% identity in
                173 aa overlap"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein NMA0702"
                /protein_id="CAB83989.1"
                /db_xref="GI:7379427"
                /translation="MSNQKALVIFSGGQDSTTCLIQAIQTYGRENVQAITFQYGQRHA
                VELERARWIAQDLGVKQTVLDLSLMRQITHNALMDDTAAIETAENGVPNTFVDGRNAL
                FLLYAAIYAKGQGIRHIIAGVCETDFSGYPDCRDVFVKSMNVTLNLAMDYDFQIHTPL
                MYLTKAQTWALADEMGALDYIREQTHTCYNGIVGGCRECPSCILRERGLAEYLESKKA
                V"
misc_feature    3522. .3531
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
RBS             3547. .3550
gene            3558. .4076
                /gene="NMA0703"
CDS             3558. .4076
                /gene="NMA0703"
                /note="NMA0703, len: 172 aa; unknown"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein NMA0703"
                /protein_id="CAB83990.1"
                /db_xref="GI:7379428"
                /translation="MPKLHMFYLGGNAGRSNIEVHDIQFAVCDDYREAVPALKAAWFG
                DADKIHIDGWQVVEWADGYDIAVSETPKTKMPSENAPRLYFANVGGYRAGQLAEAHAF
                GLFAAATPAEAKQKALQTLLTDSYVQQHKDNLKDVDNLLALDRIGNFHIRLTPNPHGK
                PAEIGFQGYLPI"
misc_feature    3782. .3791
                /gene="NMA0703"
                /note="Core DNA uptake sequence: gccgtctgaa"
                /label=DUS
misc_feature    4083. .4502
                /gene="NMA0704"
                /note="Pfam match to entry PF01242 PTPS, 6-pyruvoyl
                tetrahydropterin synthase, score 10.80, E-value 1.7e-06"
gene            4083. .4505
                /gene="NMA0704"
CDS             4083. .4505
                /gene="NMA0704"
                /note="NMA0704, len: 140 aa; similar to many hypothetical
                proteins e.g. SW:YB90_HAEIN (EMBL:U32798), HI1190,
                Haemophilus influenzae hypothetical protein (141 aa),
                fasta scores; E(): 5.1e-32, 58.0% identity in 138 aa
                overlap. Shows very weak similarity to eukaryotic
                6-pyruvoyl-tetrahydropterin synthases e.g. SW:PTPS_RAT
                (EMBL:M77850), pts, Rattus norvegicus 6-pyruvoyl
                tetrahydrobiopterin synthase precursor (EC 4.6.1.10) (144
                aa), fasta scores; E(): 0.0062, 27.3% identity in 154 aa
                overlap. Contains Pfam match to entry PF01242 PTPS,
                6-pyruvoyl tetrahydropterin synthase"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein NMA0704"
                /protein_id="CAB83991.1"

Query Match             94.3%;  Score 1315.8;  DB 2;  Length 311321;
Best Local Similarity   96.9%;  Pred. No. 8.5e-234;
Matches 1352;  Conservative    0;  Mismatches   42;  Indels    1;  Gaps    1;
```

```
Qy      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          |||||||||||||||||||| |||||||||| |||||||||||||||| ||||||||||
Db  10821 GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCCTCGCTGGCAGGC 10762

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaaagaagcatccttcgtagaacgc 120
          ||||| || || ||||||||||||||||||||||||||||||||||||||| |||||||||
Db  10761 TGCGACAAAGCCGGCAGCTTTTTCGGTGCGGACAAAAAAGAAGCATCCTTTGTAGAACGC 10702

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          ||| |||||||||||||||||||||| |||||||||||||||||||||||| |||||||
Db  10701 ATCAAACACACCAAAGACGACGGCAGCGTCAGTATGCTGCTGCCCGACTTTGTCCAACTG 10642

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10641 GTTCAAAGCGAAGGCCCGGCAGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGCACCCAA 10582

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10581 AACGGCAGCAGCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGACCCGTTCTAC 10522

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatccccccaagaagaagcagatgac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10521 GAATTTTTCAAACGCCTCGTCCCGAACATGCCCGAAATCCCCCAAGAAGAAGCAGATGAC 10462

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          |||||||||||||||||||||||||||||||||||||||| ||||||| || |||||||||
Db  10461 GGCGGATTGAACTTCGGTTCGGGCTTCATCATCAGCAAAGACGGCTATATTCTGACCAAT 10402

Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          || |||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db  10401 ACGCACGTCGTTACCGGCATGGGCAGTATCAAAGTCCTGCTCAACGACAAGCGCGAATAT 10342

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10341 ACCGCCAAACTCATCGGTTCGGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCA 10282

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          |||||||||| |||||||||||||||||||||||||| |||||||||||||||||||||
Db  10281 ACGGAAGAGCTGCCCGTCGTCAAAATCGGCAATCCCAAAGATTTGAAACCGGGCGAATGG 10222

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10221 GTCGCCGCCATCGGCGCGCCCTTCGGCTTCGACAACAGCGTGACCGCCGGCATCGTGTCC 10162

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10161 GCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCC 10102

Qy    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  10101 ATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACTTAAAAGGACAGGTCGTCGGCATC 10042

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctccttgccatcccgatt 840
          || |||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db  10041 AACTCGCAAATATACAGCCGCAGCGGCGGATTCATGGGCATTTCCTTCGCCATCCCGATT 9982

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9981 GACGTTGCCATGAATGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAA 9922

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          |||||||||||||||| |||||||||||||||||||| ||||||||| |||| |||
Db   9921 CTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTCGGTTTGGACAAA 9862

Qy    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
          ||| |||||||| |||||||||||||||| |||||||||||||||||||||||||||||
Db   9861 GCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCCCCGCAGAACGTGCCGGCCTG 9802

Qy   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9801 CGGGCGGGCGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCCGGCGACCTT 9742
```

```
Qy   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9741 CCCGTTATGGTCGGCGCCATTACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGGCGCAAA 9682

Qy   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
          |||||||||||||||||||| |||||||||||||||||||||||||| ||||||||||
Db   9681 GGCGAAGAAATCACAATCAAAGTCAAGCTGGGCAACGCCGCCGAGCATATCGGCGCATCA 9622

Qy   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9621 TCCAAAACAGATGAAGCCCCCTACACCGAACAGCAATCCGGTACGTTCTCGGTCGAATCC 9562

Qy   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db   9561 GCAGGCATTACCCTTCAGACACATACCGACAGCAGCGGCGGACACCTCGTCGTCGTACGG 9502

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          |||||||||||||||||||||||||||||| |||| |||||||||||| || |||||| ||
Db   9501 GTTTCCGACGCGGCAGAACGCGCAGGCTTGAGGCGCGGCGACGAAATTCTTGCCGTC-GG 9443

Qy   1381 gcaagtccccgtcaa 1395
          |||||||||||||||
Db   9442 GCAAGTCCCCGTCAA 9428


RESULT   4
AX011574
LOCUS       AX011574     1500 bp    DNA             BCT       18-SEP-2000
DEFINITION  Sequence 3 from Patent WO9955872.
ACCESSION   AX011574
VERSION     AX011574.1  GI:9998106
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1500)
  AUTHORS   Ruelle,J.L.
  TITLE     Basb013 dna and proteins from neisseria meningitidis
  JOURNAL   Patent: WO 9955872-A 04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
     source          1. .1500
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      370 a    469 c    392 g    269 t
ORIGIN


  Query Match             90.5%;  Score 1263;  DB 2;  Length 1500;
  Best Local Similarity   94.6%;  Pred. No. 1.5e-223;
  Matches 1319;  Conservative    0;  Mismatches   75;  Indels    1;  Gaps    1;

Qy      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          |||||||||||||||||| |||||||||| |||||||||||||||| ||||||||||
Db      1 GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCCTCGCTGGCAGGC 60

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
          ||||| || || ||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGCGACAAAGCCGGCAGCTTTTTCGGTGCGGACAAAAAAGAAGCATCCTTCGTAGAACGC 120

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          |||||||||||||||||||||||| |||||||||||||||||||||||| |||||||
Db    121 ATCGAACACACCAAAGACGACGGCAGCGTCAGTATGCTGCTGCCCGACTTTGTCCAACTG 180

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgcccgcgcacccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTTCAAAGCGAAGGCCCGGCAGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGCACCCAA 240

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AACGGCAGCGGCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGACCCGTTCTAC 300

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    301 GAATTTTTCAAACGCCTCGTCCCGAATATGCCCGAAATCCCCCAAGAAGAAGCAGATGAC 360

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          |||||||||||||||||||||||||||||||||||||||| ||||||| || |||||||||
Db    361 GGCGGATTGAACTTCGGTTCGGGCTTCATCATCAGCAAAGACGGCTATATTCTGACCAAT 420

Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          || |||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db    421 ACGCACGTCGTTACCGGCATGGGCAGTATCAAAGTCCTGCTCAACGACAAGCGCGAATAT 480

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgccccttctgaaaatcgacgca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCGCCAAACTCATCGGTTCGGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCA 540

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||
Db    541 ACGGAAGAGCTGCCCGTCGTCAAAATCGGCAATCCCAAAGATTTGAAACCGGGCGAATGG 600

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTCGCCGCCATCGGCGCGCCCTTCGGCTTCGACAACAGCGTGACCGCCGGCATCGTGTCC 660

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCC 720

Qy    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          |||||||||||||| |||||||||||||||||||| | ||||||||||||||||||||
Db    721 ATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACCTGAAAGGACAGGTCGTCGGCATC 780

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
          || ||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db    781 AACTCGCAAATATACAGCCGCAGCGGCGGATTCATGGGCATTTCCTTCGCCATCCCGATT 840

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACGTTGCCATGAATGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAA 900

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          |||||||||||||||| |||||||||||||||||||||| |||||||| |||| |||
Db    901 CTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTCGGTTTGGACAAA 960

Qy    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
          ||| |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db    961 GCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCCCCGCAGAACGTGCCGGCCTG 1020

Qy   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAGGCGGGCGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCCGGCGACCTT 1080

Qy   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCCGTTATGGTCGGCGCCATTACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGGCGCAAA 1140

Qy   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
          ||  | |||||||| |    | ||| |||||||| ||  |||| || |||||  | ||
Db   1141 GGTAAGGAAATCACCGTTGCCGTCAAACTGGGCAATGCTTCCGAACAAACCGGTTCCTCG 1200

Qy   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
          ||| |  | ||  ||||||| ||  ||||||| ||||||||||||||||||||||||||
Db   1201 TCCGAGCCGGACAAAGCCCCTTATGCCGAACACCAATCCGGTACGTTCTCGGTCGAATCC 1260

Qy   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
          |||||||||||||||||||||||||||||||||||||||||  ||  || |||||||| |||
Db   1261 GCAGGCATTACCCTTCAGACACATACCGACAGCAGCGGCGGACGGCTTGTCGTCGTGCGG 1320

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          ||||| |  |||||||||||||||||||| |||| |||||||||||||| |||||| ||
Db   1321 GTTTCGGGGGCGGCAGAACGCGCAGGCTTGAGGCGCGGCGACGAAATCCTTGCCGTC-GG 1379

Qy   1381 gcaagtccccgtcaa 1395
```

```
          ||||||||||||||||
Db   1380 GCAAGTCCCCGTCAA 1394


RESULT   5
AX011572
LOCUS       AX011572     1500 bp    DNA             BCT       18-SEP-2000
DEFINITION  Sequence 1 from Patent WO9955872.
ACCESSION   AX011572
VERSION     AX011572.1  GI:9998105
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1500)
  AUTHORS   Ruelle,J.L.
  TITLE     Basb013 dna and proteins from neisseria meningitidis
  JOURNAL   Patent: WO 9955872-A 04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
     source          1. .1500
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      370 a    468 c    393 g    269 t
ORIGIN


  Query Match           90.2%;  Score 1258.2;  DB 2;  Length 1500;
  Best Local Similarity  94.3%;  Pred. No. 1.1e-222;
  Matches 1316;  Conservative    0;  Mismatches   78;  Indels    1;  Gaps    1;

Qy       1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
           |||||||||||||||||||| |||||||||| ||||||||||||||| ||||||||||
Db       1 GTGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCCTCGCTGGCAGGC 60

Qy      61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
           ||||| || || ||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGCGACAAAGCCGGCAGCTTTTTCGGTGCGGACAAAAAAGAAGCATCCTTCGTAGAACGC 120

Qy     121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
           |||||||||||||||||||||||||| |||||||||||||||||||||||| |||||||
Db     121 ATCGAACACACCAAAGACGACGGCAGCGTCAGTATGCTGCTGCCCGACTTTGTCCAACTG 180

Qy     181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GTTCAAAGCGAAGGCCCGGCAGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGCACCCAA 240

Qy     241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AACGGCAGCGGCAATGCCGAAACCGATTCCGACCCGCTTGCCGACAGCGACCCGTTCTAC 300

Qy     301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     301 GAATTTTTCAAACGCCTCGTCCCGAATATGCCCGAAATCCCCCAAGAAGAAGCAGATGAC 360

Qy     361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
           |||||||||||||||||||||||||||||||||||||| ||||||| || ||||||||||
Db     361 GGCGGATTGAACTTCGGTTCGGGCTTCATCATCAGCAAAGACGGCTATATTCTGACCAAT 420

Qy     421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
           || ||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db     421 ACGCACGTCGTTACCGGCATGGGCAGTATCAAAGTCCTGCTCAACGACAAGCGCGAATAT 480

Qy     481 accgccaaactcatcggttcggatgtccaatccgatgtcgccccttctgaaaatcgacgca 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ACCGCCAAACTCATCGGTTCGGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCA 540

Qy     541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
           |||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||||
Db     541 ACGGAAGAGCTGCCCGTCGTCAAAATCGGCAATCCCAAAGATTTGAAACCGGGCGAATGG 600

Qy     601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
           ||||| |||||||| |||||||||||| |||||||||||||||||||||||||||||||
Db     601 GTCGCCGCCATCGGGCGGCCCTTCGGCTTCGACAACAGCGTGACCGCCGGCATCGTGTCC 660

Qy     661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GCCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCC 720

Qy     721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
           |||||||||||||| |||||||||||||||||||| | ||||||||||||||||||||
Db     721 ATCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACCTGAAAGGACAGGTCGTCGGCATC 780

Qy     781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
           || |||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db     781 AACTCGCAAATATACAGCCGCAGCGGCGGATTCATGGGCATTTCCTTCGCCATCCCGATT 840

Qy     841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GACGTTGCCATGAATGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAA 900

Qy     901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
           |||||||||||||||| |||||||||||||||||||| |||||||| |||| |||
Db     901 CTGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTCGGTTTGGACAAA 960

Qy     961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
           ||| |||||||| |||||||||||||| ||||||||||||||||||||||||||||||||
Db     961 GCCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCCCCGCAGAACGTGCCGGCCTG 1020

Qy    1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CAGGCGGGCGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCCGGCGACCTT 1080

Qy    1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CCCGTTATGGTCGGCGCCATTACGCCGGGAAAAGAAGTCAGCCTCGGCGTATGGCGCAAA 1140

Qy    1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
           ||  | |||||||  |    | ||| |||||||| ||  |||| || |||||  | ||
Db    1141 GGTAAGGAAATCACCGTTGCCGTCAAACTGGGCAATGCTTCCGAACAAACCGGTTCCTCG 1200

Qy    1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
           ||| |  | ||  |||||||  ||  |||||||  ||||||||||||||||||||||||
Db    1201 TCCGAGCCGGACAAAGCCCCTTATGCCGAACACCAATCCGGTACGTTCTCGGTCGAATCC 1260

Qy    1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
           |||||||||||||||||||||||||||||||||||||  ||  || |||||||| |||
Db    1261 GCAGGCATTACCCTTCAGACACATACCGACAGCAGCGGCGGACGGCTTGTCGTCGTGCGG 1320

Qy    1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
           ||||| |  |||||||||||||||||||| |||| |||||||||||||| |||||| ||
Db    1321 GTTTCGGGGGCGGCAGAACGCGCAGGCTTGAGGCGCGGCGACGAAATCCTTGCCGTC-GG 1379

Qy    1381 gcaagtccccgtcaa 1395
           |||||||||||||||
Db    1380 GCAAGTCCCCGTCAA 1394


RESULT   6
AX011578
LOCUS       AX011578     1110 bp    DNA             BCT       18-SEP-2000
DEFINITION  Sequence 7 from Patent WO9955872.
ACCESSION   AX011578
VERSION     AX011578.1  GI:9998108
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 1110)
  AUTHORS   Ruelle,J.L.
  TITLE     Basb013 dna and proteins from neisseria meningitidis
  JOURNAL   Patent: WO 9955872-A 04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
```

```
     source          1. .1110
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      280 a    350 c    279 g    201 t
ORIGIN


  Query Match            76.1%;  Score 1061;  DB 2;  Length 1110;
  Best Local Similarity  97.3%;  Pred. No. 2.9e-186;
  Matches 1079;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy      2 tgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggct 61
          |||||||||||||||||||| |||||||||| |||||||||||||||| ||||||||||
Db      2 TGTTCAAAAAATACCAATACCTCGCTTTGGCAGCACTGTGTGCCGCCTCGCTGGCAGGCT 61

Qy     62 gcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgca 121
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 GCGACAAGGCAGGCAGCTTTTTCGGTGCGGACAAAAAAGAAGCATCCTTCGTAGAACGCA 121

Qy    122 tcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactgg 181
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    122 TCGAACACACCAAAGACGACGGCAGCGTCAGTATGCTGCTGCCCGACTTTGCCCAACTGG 181

Qy    182 ttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaaa 241
          ||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    182 TTCAAAGTGAAGGTCCGGCAGTCGTCAATATTCAGGCAGCCCCCGCCCCGCGCACCCAAA 241

Qy    242 acggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctacg 301
          |||||||||||||||||||| |||||||||||||| ||||||||| ||||||||||||||
Db    242 ACGGCAGCGGCAATGCCGAAAACGATTCCGACCCGATTGCCGACAACGACCCGTTCTACG 301

Qy    302 aatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgacg 361
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    302 AATTTTTCAAACGCCTCGTCCCGAATATGCCCGAAATCCCCCAAGAAGAAGCAGATGACG 361

Qy    362 gcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaata 421
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    362 GCGGATTGAACTTCGGTTCGGGCTTCATCATCAGCAAAGACGGCTACATCCTGACCAATA 421

Qy    422 cccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatata 481
          |||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db    422 CCCACGTCGTTACCGGCATGGGCAGTATCAAAGTCCTGCTCAACGACAAGCGCGAATATA 481

Qy    482 ccgccaaactcatcggttcggatgtccaatccgatgtcgccccttctgaaaatcgacgcaa 541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 CCGCCAAACTCATCGGTTCGGATGTCCAATCCGATGTCGCCCTTCTGAAAATCGACGCAA 541

Qy    542 cggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatggg 601
          |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||
Db    542 CGGAAGAGCTGCCCGTCGTCAAAATCGGCAATCCCAAAGATTTGAAACCGGGCGAATGGG 601

Qy    602 tcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccg 661
          |||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    602 TCGCCGCCATCGGCGCGCCCTTCGGCTTCGACAACAGCGTGACCGCCGGCATCGTGTCCG 661

Qy    662 ccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcca 721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 CCAAAGGCAGAAGCCTGCCCAACGAAAGCTACACACCCTTCATCCAAACCGACGTTGCCA 721

Qy    722 tcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatca 781
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    722 TCAATCCGGGCAACTCCGGCGGCCCGCTGTTCAACTTAAAAGGACAGGTCGTCGGCATCA 781

Qy    782 attcgcaaatatacagccgcagcggcggattcatgggcatctccttgccatcccgattg 841
          | |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||
Db    782 ACTCGCAAATATACAGCCGCAGCGGCGGATTCATGGGCATTTCCTTCGCCATCCCGATTG 841

Qy    842 acgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaac 901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    842 ACGTTGCCATGAATGTCGCCGAACAGCTGAAAAACACCGGCAAAGTCCAACGCGGACAAC 901

Qy    902 tgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaag 961
          |||||||||||||||| |||||||||||||||||||||| ||||||||| |||| ||||

Db    902 TGGGCGTGATTATTCAAGAAGTATCCTACGGTTTGGCACAATCGTTCGGTTTGGACAAAG 961

Qy    962 ccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgc 1021
          || |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db    962 CCGGCGGCGCACTGATTGCCAAAATCCTGCCCGGCAGCCCCGCAGAACGTGCCGGCCTGC 1021

Qy   1022 aggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttc 1081
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1022 AGGCGGGCGACATCGTCCTCAGCCTCGACGGCGGAGAAATACGTTCTTCCGGCGACCTTC 1081

Qy   1082 ccgtcatggtcggcgccattacgccggga 1110
          |||| |||||||||||||||||||||||
Db   1082 CCGTTATGGTCGGCGCCATTACGCCGGGA 1110


RESULT   7
PAU32853
LOCUS       PAU32853     1961 bp    DNA             BCT       06-MAR-1996
DEFINITION  Pseudomonas aeruginosa mucC and mucD genes, complete cds.
ACCESSION   U32853
VERSION     U32853.1  GI:1184682
KEYWORDS    .
SOURCE      Pseudomonas aeruginosa.
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 1 to 1961)
  AUTHORS   Boucher,J.C., Martinez-Salazar,J., Schurr,M.J., Mudd,M.H., Yu,H.
            and Deretic,V.
  TITLE     Two distinct loci affecting conversion to mucoidy in Pseudomonas
            aeruginosa in cystic fibrosis encode homologs of the serine
            protease HtrA
  JOURNAL   J. Bacteriol. 178 (2), 511-523 (1996)
  MEDLINE   96134987
REFERENCE   2  (bases 1 to 1961)
  AUTHORS   Boucher,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-1995) John C. Boucher, Microbiology, University
            of Texas Health Science Center, 7703 Floyd Curl Dr., San Antonio,
            TX 78240, USA
FEATURES             Location/Qualifiers
     source          1. .1961
                     /organism="Pseudomonas aeruginosa"
                     /strain="pao1"
                     /db_xref="taxon:287"
                     /map="67.5 min"
     gene            1. .456
                     /gene="mucC"
     CDS             1. .456
                     /gene="mucC"
                     /codon_start=1
                     /transl_table=11
                     /product="MucC"
                     /protein_id="AAC43675.1"
                     /db_xref="GI:1184683"
                     /translation="MIEEQGRVVATEPGAVWVETVRRSTCSSCSANAGCGQGLMQRLG
                     VGAGRARVRALSDLSLRVGDAVVLGIHEDLLLRASVLFYLFPLLGFFVAALLATRAGL
                     VEPLIIVSGLAGLLAAWLLVRRHARRHADDPASQPVVLRALISGPSDSA"
     gene            496. .1920
                     /gene="mucD"
     CDS             496. .1920
                     /gene="mucD"
                     /note="HtrA serine protease homolog; similar to E. coli
                     HtrA, Swiss-Prot Accession Number P09376"
                     /codon_start=1
                     /transl_table=11
                     /product="MucD"
                     /protein_id="AAC43676.1"
                     /db_xref="GI:1184684"
                     /translation="MHTLKRCMAAMVALLALSLAMTARAELPDFTPLVEQASPAVVNI
                     STRQKLPDRAMARGQLSIPDLEGLPPMFRDFLERSIPQVPRNPRGQQREAQSLGSGFI
                     ISNDGYILTNNHVVADADEILVRLSDRSEHKAKLIGADPRSDVAVLKIEAKNLPTLKL
                     GDSNKLKVGEWVLAIGSPFGFDHSVTAGIVSAKGRSLPNESYVPFIQTDVAINPGNSG
```

```
                     GPLLNLQGEVVGINSQIFTRSGGFMGLSFAIPIDVALNVADQLKKAGKVSRGWLGVVI
                     QEVNKDLAESFGLDKPSGALVAQLVEDGPAAKGGLQVGDVILSLNGQSINESADLPHL
                     VGNMKPGDKINLDVIRNGQRKSLSMAVGSLPDDDEEIASMGAPGAERSSNRLGVTVAD
                     LTAEQRKSLDIQGGVVIKEVQDGPAAVIGLRPGDVITHLDNKAVTSTKVFADVAKALP
                     KNRSVSMRVLRQGRASFITFKLAE"
BASE COUNT      332 a    614 c    657 g    358 t
ORIGIN


  Query Match           26.2%;  Score 365.6;  DB 2;  Length 1961;
  Best Local Similarity  68.0%;  Pred. No. 5.1e-58;
  Matches  525;  Conservative    0;  Mismatches  244;  Indels    3;  Gaps    1;

Qy    374 tcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttg 433
          | ||||| |||||||||||||  |||  |||||||||||||| ||||| |  |||||||| |
Db    785 TGGGTTCCGGCTTCATCATCTCCAACGACGGCTACATCCTCACCAACAATCACGTCGTGG 844

Qy    434 ccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactca 493
          ||| |   | |   |||   || |  ||   ||||    ||||  | |  |||||  | |
Db    845 CCGATGCCGACGAGATCCTGGTGCGCCTGTCCGACCGTAGCGAGCACAAGGCCAAGTTGA 904

Qy    494 tcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctac 553
          ||||  ||||    |    ||| || ||  | ||||| ||||| |   ||  |  || |
Db    905 TCGGCGCGGACCCGCGCAGCGACGTGGCGGTGCTGAAGATCGAGG---CGAAGAACCTGC 961

Qy    554 ccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcg 613
          |   | | ||| | ||| || | || ||  |||||  ||||||||||||||   |||||||
Db    962 CGACCCTGAAACTGGGCGATTCGAACAAGCTGAAAGTGGGCGAATGGGTCCTGGCCATCG 1021

Qy    614 gcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcagaa 673
          |  |||| ||||||||  ||   ||   || |||||||| |||||    ||||| ||  | |
Db   1022 GTTCGCCGTTCGGCTTCGATCACTCGGTCACCGCCGGTATCGTCAGTGCCAAGGGTCGTA 1081

Qy    674 gcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggca 733
          |||||||  |||||  ||||||  |||||||||||| ||||||||  || |||||  ||||| |
Db   1082 GCCTGCCGAACGAGAGCTACGTACCCTTCATCCAGACCGACGTGGCGATCAACCCGGGTA 1141

Qy    734 attccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatat 793
          | ||||||||  ||||||  | |||  |  ||  | ||  ||||||||||| |||||  ||  |
Db   1142 ACTCCGGCGGTCCGCTGCTGAACCTGCAGGGCGAAGTGGTCGGCATCAACTCGCAGATCT 1201

Qy    794 acagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatga 853
           || |||   |||||| |||||||||  | |||||  |||||  |||||  || ||  ||  |||
Db   1202 TCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGATGTCGCGCTGA 1261

Qy    854 atgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgatta 913
          | ||||||||  |||  ||||  ||  |||||||  |||   |||||    |||||||||  | |
Db   1262 ACGTCGCCGACCAGTTGAAGAAAGCCGGCAAGGTCAGTCGCGGCTGGCTGGGCGTGGTGA 1321

Qy    914 ttcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcat 973
          | ||||||||   | | | | | ||  ||||  |||||  || ||  ||  |   ||||||
Db   1322 TCCAGGAAGTGAACAAGGATCTCGCCGAGTCCTTCGGCCTCGACAAGCCGTCCGGCGCGC 1381

Qy    974 tgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgaca 1033
          || | ||  || |  |    | |  |||| || |      |  ||||||||| ||||||
Db   1382 TGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCCAAGGGTGGCCTGCAGGTGGGCGATG 1441

Qy   1034 tcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcg 1093
          |  ||||||||||  |||||      ||       |||| |||||| ||   | |||| |
Db   1442 TGATCCTCAGCCTGAACGGCCAGTCGATCAACGAGTCCGCCGACCTGCCGCACCTGGTGG 1501

Qy   1094 gcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcga 1145
          ||  ||| | ||||||  |  |  ||| ||| | |||    ||||| ||| |
Db   1502 GCAACATGAAGCCGGGCGACAAGATCAACCTGGACGTGATTCGCAACGGCCA 1553


RESULT   8
PAU49151
LOCUS       PAU49151     4587 bp    DNA             BCT       13-MAR-1996
DEFINITION  Pseudomonas aeruginosa alternate sigma factor (algU), mucA, mucB,
            mucC and mucD genes, complete cds.
ACCESSION   U49151 U24569 U08380 L04794 L02119
VERSION     U49151.1  GI:1220191
KEYWORDS    .
SOURCE      Pseudomonas aeruginosa strain~PAO1.
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 409 to 1055)
  AUTHORS   Martin,D.W., Holloway,B.W. and Deretic,V.
  TITLE     Characterization of a locus determining the mucoid status of
            Pseudomonas aeruginosa: AlgU shows sequence similarities with a
            Bacillus sigma factor
  JOURNAL   J. Bacteriol. 175 (4), 1153-1164 (1993)
  MEDLINE   93163045
REFERENCE   2  (bases 528 to 2836)
  AUTHORS   Martin,D.W., Schurr,M.J., Mudd,M.H. and Deretic,V.
  TITLE     Differentiation of Pseudomonas aeruginosa into the
            alginate-producing form: inactivation of mucB causes conversion to
            mucoidy
  JOURNAL   Mol. Microbiol. 9 (3), 497-506 (1993)
  MEDLINE   94018645
REFERENCE   3  (sites)
  AUTHORS   Martin,D.W., Schurr,M.J., Mudd,M.H., Govan,J.R., Holloway,B.W. and
            Deretic,V.
  TITLE     Mechanism of conversion to mucoidy in Pseudomonas aeruginosa
            infecting cystic fibrosis patients
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 90 (18), 8377-8381 (1993)
  MEDLINE   93391358
REFERENCE   4  (bases 252 to 488)
  AUTHORS   Martin,D.W., Schurr,M.J., Yu,H. and Deretic,V.
  TITLE     Analysis of promoters controlled by the putative sigma factor AlgU
            regulating conversion to mucoidy in Pseudomonas aeruginosa:
            relationship to sigma E and stress response
  JOURNAL   J. Bacteriol. 176 (21), 6688-6696 (1994)
  MEDLINE   95050230
REFERENCE   5  (sites)
  AUTHORS   Yu,H., Schurr,M.J. and Deretic,V.
  TITLE     Functional equivalence of Escherichia coli sigma E and Pseudomonas
            aeruginosa AlgU: E. coli rpoE restores mucoidy and reduces
            sensitivity to reactive oxygen intermediates in algU mutants of P.
            aeruginosa
  JOURNAL   J. Bacteriol. 177 (11), 3259-3268 (1995)
  MEDLINE   95286510
REFERENCE   6  (bases 1 to 252)
  AUTHORS   Schurr,M.J., Yu,H., Boucher,J.C., Hibler,N.S. and Deretic,V.
  TITLE     Multiple promoters and induction by heat shock of the gene encoding
            the alternative sigma factor AlgU (sigma E) which controls mucoidy
            in cystic fibrosis isolates of Pseudomonas aeruginosa
  JOURNAL   J. Bacteriol. 177 (19), 5670-5679 (1995)
  MEDLINE   96032406
REFERENCE   7  (bases 1 to 4587)
  AUTHORS   Boucher,J.C., Martinez-Salazar,J., Schurr,M.J., Mudd,M.H., Yu,H.
            and Deretic,V.
  TITLE     Two distinct loci affecting conversion to mucoidy in Pseudomonas
            aeruginosa in cystic fibrosis encode homologs of the serine
            protease HtrA
  JOURNAL   J. Bacteriol. 178 (2), 511-523 (1996)
  MEDLINE   96134987
REFERENCE   8  (bases 1 to 4587)
  AUTHORS   Boucher,J.C., Martin,D.W., Schurr,M.J., Deretic,V., Yu,H., Mudd,M
            and Martinez-Salizar,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1996) John C. Boucher, Microbiology, U.T. HSC San
            Antonio, 7703 Floyd Curl Dr., San Antonio, Tx 78284, USA
COMMENT     On Mar 11, 1996 this sequence version replaced gi:1173501.
FEATURES             Location/Qualifiers
     source          1. .4587
                     /organism~"Pseudomonas aeruginosa"
                     /strain~"PAO1"
                     /db_xref~"taxon:287"
                     /map="67.5"
     mRNA            22
                     /note~"transcription start site for promoter P5; 3' end is
                     unknown"
     mRNA            124
```

```
                     /note="transcription start site for promoter P4; 3' end is
                     unknown"
     promoter        204. .238
                     /note="P3 promoter"
     -35_signal      204. .209
                     /note="P3 promoter"
     -10_signal      226. .230
                     /note="P3 promoter"
     mRNA            235
                     /note="AlgU-dependent transcription start site for
                     promoter P3; 3' end is unknown"
     mRNA            238
                     /note="AlgU-dependent transcription start site for
                     promoter P3; 3' end is unknown"
     mRNA            310
                     /note="transcription start site for promoter P2; 3' end is
                     unknown"
     -35_signal      389. .394
                     /note="P1 promoter"
     promoter        389. .421
                     /note="P1 promoter, the first of two AlgU-dependent
                     promoters; AlgU-dependent promoters strongly resemble the
                     sigmaE promoters of E. coli in both the -10 and -35
                     sequences"
     -10_signal      411. .415
                     /note="P1 promoter"
     mRNA            421
                     /note="transcription start site for promoter P1; 3' end is
                     unknown"
     gene            474. .1055
                     /gene="algU"
     CDS             474. .1055
                     /gene="algU"
                     /codon_start=1
                     /transl_table=11
                     /function="conversion to mucoidy;  required for algD
                     transcription"
                     /product="sigma factor"
                     /protein_id="AAC43714.1"
                     /db_xref="GI:1173502"
                     /translation="MLTQEQDQQLVERVQRGDKRAFDLLVLKYQHKILGLIVRFVHDA
                     QEAQDVAQEAFIKAYRALGNFRGDSAFYTWLYRIAINTAKNHLVARGRRPPDSDVTAE
                     DAEFFEGDHALKDIESPERAMLRDEIEATVHQTIQQLPEDLRTALTLREFEGLSYEDI
                     ATVMQCPVGTVRSRIFRAREAIDKALQPLLREA"
     gene            1087. .1671
                     /gene="mucA"
     CDS             1087. .1671
                     /gene="mucA"
                     /codon_start=1
                     /transl_table=11
                     /product="MucA"
                     /protein_id="AAC43715.1"
                     /db_xref="GI:1173503"
                     /translation="MSREALQETLSAVMDNEADELELRRVLAACGEDAELRSTWSRYQ
                     LARSVMHREPTLPKLDIAAAVSAALADEAAPPKAEKGPWRMVGRLAVAASVTLAVLAG
                     VRLYNQNDALPQMAQQGTTPQIALPQVKGPAVLAGYSEEQGAPQVITNSSSSDTRWHE
                     QRLPIYLRQHVQQSAVSGTESALPYARAASLENR"
     gene            1680. .2630
                     /gene="mucB"
     CDS             1680. .2630
                     /gene="mucB"
                     /codon_start=1
                     /transl_table=11
                     /product="MucB"
                     /protein_id="AAC43716.1"
                     /db_xref="GI:1220192"
                     /translation="MRTTSLLLLLGSLMAVPATQAADASDWLNRLAEADRQNSFQGTF
                     VYERNGSFSTHEIWHRVESDGAVRERLLQLDGARQEVVRVDGRTQCISGGLADQLADA
                     QLWPVRKFDPSQLASWYDLRLVGESRVAGRPAVVLAVTPRDQHRYGFELHLDRDTGLP
                     LKSLLLNEKGQLLERFQFTQLNTGAAPAEDQLQAGAECQVVGPAKADGEKTVAWRSEW
                     LPPGFTLTRSFMRRSPVTPDPVACLTYGDGLARFSVFIEPLHGAMVGDARSQLGPTVV
                     VSKRLQTDDGGQMVTVVGEVPLGTAERVALSIRPEAAAQK"
     gene            2627. .3082
                     /gene="mucC"
     CDS             2627. .3082
                     /gene="mucC"
                     /codon_start=1
                     /transl_table=11
                     /product="MucC"
                     /protein_id="AAC43717.1"
                     /db_xref="GI:1220193"
                     /translation="MIEEQGRVVATEPGAVWVETVRRSTCSSCSANAGCGQGLMQRLG
                     VGAGRARVRALSDLSLRVGDAVVLGIHEDLLLRASVLFYLFPLLGFFVAALLATRAGL
                     VEPLIIVSGLAGLLAAWLLVRRHARRHADDPASQPVVLRALISGPSDSA"
     gene            3122. .4546
                     /gene="mucD"
     CDS             3122. .4546
                     /gene="mucD"
                     /codon_start=1
                     /transl_table=11
                     /product="MucD"
                     /protein_id="AAC43718.1"
                     /db_xref="GI:1220194"
                     /translation="MHTLKRCMAAMVALLALSLAMTARAELPDFTPLVEQASPAVVNI
                     STRQKLPDRAMARGQLSIPDLEGLPPMFRDFLERSIPQVPRNPRGQQREAQSLGSGFI
                     ISNDGYILTNNHVVADADEILVRLSDRSEHKAKLIGADPRSDVAVLKIEAKNLPTLKL
                     GDSNKLKVGEWVLAIGSPFGFDHSVTAGIVSAKGRSLPNESYVPFIQTDVAINPGNSG
                     GPLLNLQGEVVGINSQIFTRSGGFMGLSFAIPIDVALNVADQLKKAGKVSRGWLGVVI
                     QEVNKDLAESFGLDKPSGALVAQLVEDGPAAKGGLQVGDVILSLNGQSINESADLPHL
                     VGNMKPGDKINLDVIRNGQRKSLSMAVGSLPDDDEEIASMGAPGAERSSNRLGVTVAD
                     LTAEQRKSLDIQGGVVIKEVQDGPAAVIGLRPGDVITHLDNKAVTSTKVFADVAKALP
                     KNRSVSMRVLRQGRASFITFKLAE"
BASE COUNT      811 a   1447 c   1515 g    814 t
ORIGIN
```

```
  Query Match             26.2%;  Score 365.6;  DB 2;  Length 4587;
  Best Local Similarity   68.0%;  Pred. No. 4.3e-58;
  Matches  525;  Conservative    0;  Mismatches  244;  Indels    3;  Gaps    1;

Qy    374 tcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttg 433
          | ||||| |||||||||||  |||  |||||||||||||| ||||| |  |||||||| |
Db   3411 TGGGTTCCGGCTTCATCATCTCCAACGACGGCTACATCCTCACCAACAATCACGTCGTGG 3470

Qy    434 ccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactca 493
          ||| |   | |   |||   || |  ||   ||||    |||| | |  ||||  | |
Db   3471 CCGATGCCGACGAGATCCTGGTGCGCCTGTCCGACCGTAGCGAGCACAAGGCCAAGTTGA 3530

Qy    494 tcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctac 553
          ||||  ||||    |    ||| || ||  | ||||| ||||| |   || |  | || |
Db   3531 TCGGCGCGGACCCGCGCAGCGACGTGGCGGTGCTGAAGATCGAGG---CGAAGAACCTGC 3587

Qy    554 ccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcg 613
          |   | | ||| | ||| || | || ||  |||||  ||||||||||||||    |||||||
Db   3588 CGACCCTGAAACTGGGCGATTCGAACAAGCTGAAAGTGGGCGAATGGGTCCTGGCCATCG 3647

Qy    614 gcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcagaa 673
          |  |||| |||||||| ||  ||   || |||||||| |||||    ||||| ||  | |
Db   3648 GTTCGCCGTTCGGCTTCGATCACTCGGTCACCGCCGGTATCGTCAGTGCCAAGGGTCGTA 3707

Qy    674 gcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggca 733
          ||||||| ||||| ||||||  |||||||||||| |||||||| || ||||| ||||| |
Db   3708 GCCTGCCGAACGAGAGCTACGTACCCTTCATCCAGACCGACGTGGCGATCAACCCGGGTA 3767

Qy    734 attccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatat 793
          | |||||||| |||||| | ||| |  | ||  | || ||||||||||| ||||| || |
Db   3768 ACTCCGGCGGTCCGCTGCTGAACCTGCAGGGCGAAGTGGTCGGCATCAACTCGCAGATCT 3827

Qy    794 acagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatga 853
            || |||   |||||| |||||||||| | ||||| ||||| ||||| || || ||  |||
Db   3828 TCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGATGTCGCGCTGA 3887

Qy    854 atgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgatta 913
          | |||||||| ||| |||| ||  ||||||| |||   |||||    |||||||||| | |
Db   3888 ACGTCGCCGACCAGTTGAAGAAAGCCGGCAAGGTCAGTCGCGGCTGGCTGGGCGTGGTGA 3947
```

```
Qy     914 ttcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcat 973
           | ||||||||   | | | | | ||  |||| ||||| || || ||  |   ||||||
Db    3948 TCCAGGAAGTGAACAAGGATCTCGCCGAGTCCTTCGGCCTCGACAAGCCGTCCGGCGCGC 4007

Qy     974 tgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgaca 1033
           || | ||  || |  |    | | |||| || |     |  |||||||||| ||||||
Db    4008 TGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCCAAGGGTGGCCTGCAGGTGGGCGATG 4067

Qy    1034 tcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcg 1093
           |  |||||||||  |||||      ||       |||| ||||| ||   | |||| |
Db    4068 TGATCCTCAGCCTGAACGGCCAGTCGATCAACGAGTCCGCCGACCTGCCGCACCTGGTGG 4127

Qy    1094 gcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcga 1145
           ||  ||| | ||||||  |  |  ||| ||| | |||    ||||| ||| |
Db    4128 GCAACATGAAGCCGGGCGACAAGATCAACCTGGACGTGATTCGCAACGGCCA 4179


RESULT   9
AE004511
LOCUS       AE004511    10977 bp    DNA             BCT       30-AUG-2000
DEFINITION  Pseudomonas aeruginosa PA01, section 72 of 529 of the complete
            genome.
ACCESSION   AE004511 AE004091
VERSION     AE004511.1  GI:9946646
KEYWORDS    .
SOURCE      Pseudomonas aeruginosa.
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 1 to 10977)
  AUTHORS   Stover,C.K., Pham,X.Q., Erwin,A.L., Mizoguchi,S.D., Warrener,P.,
            Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M.,
            Garber,R.L., Goltry,L., Tolentino,E., Westbrock-Wadman,S., Yuan,Y.,
            Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R.,
            Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T.
  TITLE     Complete genome sequence of Pseudomonas aeruginosa PA01, an
            opportunistic pathogen
  JOURNAL   Nature 406 (6799), 959-964 (2000)
  MEDLINE   20437337
REFERENCE   2  (bases 1 to 10977)
  AUTHORS   Stover,C.K., Pham,X.-Q.T., Erwin,A.L., Mizoguchi,S.D., Warrener,P.,
            Hickey,M.J., Brinkman,F.S.L., Hufnagle,W.O., Kowalik,D.J.,
            Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E.,
            Westbrook-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N.,
            Folger,K.R., Kas,A., Larbig,K., Lim,R.M., Smith,K.A., Spencer,D.H.,
            Wong,G.K.-S., Wu,Z., Paulsen,I.T., Reizer,J., Saier,M.H.,
            Hancock,R.E.W., Lory,S. and Olson,M.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2000) Department of Medicine and Genetics,
            University of Washington Genome Center, University Of Washington,
            Box 352145, Seattle, WA 98195, USA
FEATURES             Location/Qualifiers
     source          1. .10977
                     /organism="Pseudomonas aeruginosa"
                     /strain="PAO1"
                     /db_xref="taxon:287"
     gene            complement(81. .1025)
                     /gene="PA0759"
     CDS             complement(81. .1025)
                     /gene="PA0759"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAG04148.1"
                     /db_xref="GI:9946647"
                     /translation="MADPAFHTPLVHEGILAVRGPDAAKFLQGQLTCNLAYLNDETSS
                     LGGRCNIKGRLLSSFRILPEGDGLLLAMAGELLEAQLADLKKYAVFSKASLADESAAW
                     LRIGLRDASEALRALGIDTPAESGRIARHGDLLAVALGDGRVELWVPAQRAEAVLATL
                     REHSREAPLDDWLLGQVRAGIGQVFGATRELFIPQMINLQAVGGVSFKKGCYTGQEIV
                     ARMQYLGRLKRRLYRLALDADEVPAPGTGLFSPVHSTSVGEVVLAARTPENSVELLAV
                     LQDDAVADGRISLGSAEGAPLVLLNLPYTLDSDREIQR"
     gene            1298. .1552
                     /gene="PA0760"
     CDS             1298. .1552
                     /gene="PA0760"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAG04149.1"
                     /db_xref="GI:9946648"
                     /translation="MADETELKRLYWHSRRGMLELDVLLVPFVQEVYPGLDAEDQARF
                     RKLLECEDQDMFGWFMQRGEPEDADLRRMVRMILDRVQPD"
     gene            complement(1957. .3573)
                     /gene="nadB"
                     /note="PA0761"
     CDS             complement(1957. .3573)
                     /gene="nadB"
                     /codon_start=1
                     /transl_table=11
                     /product="L-aspartate oxidase"
                     /protein_id="AAG04150.1"
                     /db_xref="GI:9946649"
                     /translation="MSQHYQHDVLVIGSGAAGLSLALTLPEHLRIAVLSKGELSQGST
                     YWAQGGVAAVLDDTDTVESHVEDTLVAGGGLCREDAVRFTVEHSREAIQWLIEQGVPF
                     TRDEQHNHEEGSFEYHLTREGGHSHRRIIHAADATGAAIFNTLLAQARRRPNIELLSQ
                     RVAVDLITERKLGLPSKRCLGAYVLNRESGEVDTFRARFSVLASGGASKVYLYTSNPD
                     GNSGDGIAMAWRAGCRVGNLEFNQFHPTCLYHPQAKSFLITEALRGEGALLRLPNGER
                     FMPRFDPRGELAPRDIVARAIDHEMKRLGIDCVYLDISHKPAEFIKAHFPTVYERCLD
                     FGIDITQQPIPVVPAAHYTCGGVLVDQHGHTDVPGLYAIGETTFTGLHGANRMASNSL
                     LECFVYARSAAADMLQRLPGTPVPESLPSWDASQVTDSDEDVIIAHNWDELRRFMWDY
                     VGIVRTNKRLQRAQHRVRLLLSEIDEFYSNYKVSRDLIELRNLALVAELIIRSAMQRR
                     ESRGLHYTLDYPDLLPEARDTILVPPIYGD"
     gene            3982. .4563
                     /gene="algU"
                     /note="PA0762"
     CDS             3982. .4563
                     /gene="algU"
                     /codon_start=1
                     /transl_table=11
                     /product="sigma factor AlgU"
                     /protein_id="AAG04151.1"
                     /db_xref="GI:9946650"
                     /translation="MLTQEQDQQLVERVQRGDKRAFDLLVLKYQHKILGLIVRFVHDA
                     QEAQDVAQEAFIKAYRALGNFRGDSAFYTWLYRIAINTAKNHLVARGRRPPDSDVTAE
                     DAEFFEGDHALKDIESPERAMLRDEIEATVHQTIQQLPEDLRTALTLREFEGLSYEDI
                     ATVMQCPVGTVRSRIFRAREAIDKALQPLLREA"
     gene            4595. .5179
                     /gene="mucA"
                     /note="PA0763"
     CDS             4595. .5179
                     /gene="mucA"
                     /codon_start=1
                     /transl_table=11
                     /product="anti-sigma factor MucA"
                     /protein_id="AAG04152.1"
                     /db_xref="GI:9946651"
                     /translation="MSREALQETLSAVMDNEADELELRRVLAACGEDAELRSTWSRYQ
                     LARSVMHREPTLPKLDIAAAVSAALADEAAPPKAEKGPWRMVGRLAVAASVTLAVLAG
                     VRLYNQNDALPQMAQQGTTPQIALPQVKGPAVLAGYSEEQGAPQVITNSSSSDTRWHE
                     QRLPIYLRQHVQQSAVSGTESALPYARAASLENR"
     gene            5188. .6138
                     /gene="mucB"
                     /note="PA0764"
     CDS             5188. .6138
                     /gene="mucB"
                     /codon_start=1
                     /transl_table=11
                     /product="negative regulator for alginate biosynthesis
                     MucB"
                     /protein_id="AAG04153.1"
                     /db_xref="GI:9946652"
                     /translation="MRTTSLLLLLGSLMAVPATQAADASDWLNRLAEADRQNSFQGTF
                     VYERNGSFSTHEIWHRVESDGAVRERLLQLDGARQEVVRVDGRTQCISGGLADQLADA
                     QLWPVRKFDPSQLASWYDLRLVGESRVAGRPAVVLAVTPRDQHRYGFELHLDRDTGLP
                     LKSLLLNEKGQLLERFQFTQLNTGAAPAEDQLQAGAECQVVGPAKADGEKTVAWRSEW
                     LPPGFTLTRSFMRRSPVTPDPVACLTYGDGLARFSVFIEPLHGAMVGDARSQLGPTVV
```

```
                VSKRLQTDDGGQMVTVVGEVPLGTAERVALSIRPEAAAQK"
     gene       6135. .6590
                /gene="mucC"
                /note="PA0765"
     CDS        6135. .6590
                /gene="mucC"
                /codon_start=1
                /transl_table=11
                /product="positive regulator for alginate biosynthesis
                MucC"
                /protein_id="AAG04154.1"
                /db_xref="GI:9946653"
                /translation="MIEEQGRVVATEPGAVWVETVRRSTCSSCSANAGCGQGLMQRLG
                VGAGRARVRALSDLSLRVGDAVVLGIHEDLLLRASVLFYLFPLLGFFVAALLATRAGL
                VEPLIIVSGLAGLLAAWLLVRRHARRHADDPASQPVVLRALISGPSDSA"
     gene       6630. .8054
                /gene="mucD"
                /note="PA0766"
     CDS        6630. .8054
                /gene="mucD"
                /codon_start=1
                /transl_table=11
                /product="serine protease MucD precursor"
                /protein_id="AAG04155.1"
                /db_xref="GI:9946654"
                /translation="MHTLKRCMAAMVALLALSLAMTARAELPDFTPLVEQASPAVVNI
                STRQKLPDRAMARGQLSIPDLEGLPPMFRDFLERSIPQVPRNPRGQQREAQSLGSGFI
                ISNDGYILTNNHVVADADEILVRLSDRSEHKAKLIGADPRSDVAVLKIEAKNLPTLKL
                GDSNKLKVGEWVLAIGSPFGFDHSVTAGIVSAKGRSLPNESYVPFIQTDVAINPGNSG
                GPLLNLQGEVVGINSQIFTRSGGFMGLSFAIPIDVALNVADQLKKAGKVSRGWLGVVI
                QEVNKDLAESFGLDKPSGALVAQLVEDGPAAKGGLQVGDVILSLNGQSINESADLPHL
                VGNMKPGDKINLDVIRNGQRKSLSMAVGSLPDDDEEIASMGAPGAERSSNRLGVTVAD
                LTAEQRKSLDIQGGVVIKEVQDGPAAVIGLRPGDVITHLDNKAVTSTKVFADVAKALP
                KNRSVSMRVLRQGRASFITFKLAE"
     gene       8204. .10003
                /gene="le"
                /note="PA0767"
     CDS        8204. .10003
                /gene="le"
                /codon_start=1
                /transl_table=11
                /product="GTP-binding protein LepA"
                /protein_id="AAG04156.1"
                /db_xref="GI:9946655"
                /translation="MSDLSHIRNFSIIAHIDHGKSTLADRFIQMCGGLSDREMEAQVL
                DSMDLERERGITIKAHSVTLHYKAQDGKTYQLNFIDTPGHVDFTYEVSRSLAACEGAL
                LVVDAGQGVEAQSVANCYTAIEQGLEVMPVLNKMDLPQAEPERVKEEIESIIGIDATD
                AVACSAKSGMGVLEVLERLVTAIPAPEGEIEAPLQALIIDSWFDNYLGVVSLVRVKNG
                RVKKGDKILVKSTGKVHQVDSVGVFTPKHTETVDLKAGEVGFIIAGIKDIHGAPVGDT
                LTLNNTPDVEVLPGFKRVKPQVYAGLFPVSSDDFEDFRDALQKLTLNDSSLQYEPESS
                EALGFGFRCGFLGMLHMEIIQERLEREYDLDLITTAPTVVFEIVQKNGEIIYVDNPSK
                LPDLASIQEMREPICRATILVPKDHLGNVITLCIEKRGVQRDMHFLSGQVQVVYDLPM
                NEVVLDFFDRLKSTSRGYASLDYSFDRFEPSNLVRLDVLINGEKVDALALIVHRDNAP
                YKGRQLVEKMKELIPRQMFDVAIQAAIGGQIIARSTVKALRKNVLAKCYGGDVSRKRK
                LLEKQKAGKKRMKQVGSVEIPQEAFLAVLKVDS"
     gene       10009. .10863
                /gene="lepB"
                /note="PA0768"
     CDS        10009. .10863
                /gene="lepB"
                /codon_start=1
                /transl_table=11
                /product="signal peptidase I"
                /protein_id="AAG04157.1"
                /db_xref="GI:9946656"
                /translation="MTLNFPLLLVIAVAVCGALALVDLVLFAPRRRAAISSYEGQVNE
                PDPAVLEKLNKEPLLVEYGKSFFPVLFIVLVLRSFLVEPFQIPSGSMKPTLEVGDFIL
                VNKFAYGIRLPVLDTKVIPIGDPQRGDVMVFRYPSEPNINYIKRVVGLPGDTVRYTKE
                KRLYVNGELVAEKLVGEEPGTLGSVTLYQEKLGQAEHLIRKEMSRYRIEPDRQWTIPA
                GHYFMMGDNRDNSNDSRYWNDPKIPKDLLGMVPDRNIVGKAFAVWMSWPDPKMSNLPN
                FSRVGVIH"
BASE COUNT     1898 a   3459 c   3687 g   1933 t
ORIGIN
```

```
  Query Match             26.2%;  Score 365.6;  DB 1;  Length 10977;
  Best Local Similarity   68.0%;  Pred. No. 3.5e-58;
  Matches  525;  Conservative    0;  Mismatches  244;  Indels    3;  Gaps    1;

Qy     374 tcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttg 433
           | ||||| |||||||||||||  |||  ||||||||||||| ||||| | ||||||||
Db    6919 TGGGTTCCGGCTTCATCATCTCCAACGACGGCTACATCCTCACCAACAATCACGTCGTGG 6978

Qy     434 ccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactca 493
           ||| |   | |   |||   || | ||   ||||    ||||  | |  |||||  | |
Db    6979 CCGATGCCGACGAGATCCTGGTGCGCCTGTCCGACCGTAGCGAGCACAAGGCCAAGTTGA 7038

Qy     494 tcggttcggatgtccaatccgatgtcgccctttctgaaaatcgacgcaacggaagagctac 553
           ||||  ||||    |    ||| || ||  | ||||| ||||| |   || |  | || |
Db    7039 TCGGCGCGGACCCGCGCAGCGACGTGGCGGTGCTGAAGATCGAGG---CGAAGAACCTGC 7095

Qy     554 ccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcg 613
           |   | | ||| | ||| || | || ||  |||||  ||||||||||||||   |||||||
Db    7096 CGACCCTGAAACTGGGCGATTCGAACAAGCTGAAAGTGGGCGAATGGGTCCTGGCCATCG 7155

Qy     614 gcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcagaa 673
           |  |||| |||||||| ||  ||   || |||||||| |||||    |||| ||  | |
Db    7156 GTTCGCCGTTCGGCTTCGATCACTCGGTCACCGCCGGTATCGTCAGTGCCAAGGGTCGTA 7215

Qy     674 gcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggca 733
           |||||||  |||||  ||||||  |||||||||||| |||||||| || ||||| ||||| |
Db    7216 GCCTGCCGAACGAGAGCTACGTACCCTTCATCCAGACCGACGTGGCGATCAACCCGGGTA 7275

Qy     734 attccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatat 793
           | |||||||  |||||  | ||| |  | ||  | || ||||||||||  ||||| || |
Db    7276 ACTCCGGCGGTCCGCTGCTGAACCTGCAGGGCGAAGTGGTCGGCATCAACTCGCAGATCT 7335

Qy     794 acagccgcagcggcggattcatgggcatctccttttgccatcccgattgacgttgccatga 853
           || |||   ||||||  |||||||||  | |||||  |||||  |||||  || || ||  |||
Db    7336 TCACCCGTTCCGGCGGCTTCATGGGCCTGTCCTTCGCCATTCCGATCGATGTCGCGCTGA 7395

Qy     854 atgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgatta 913
           | ||||||||  ||| |||| ||  |||||||  |||   |||||    |||||||||  | |
Db    7396 ACGTCGCCGACCAGTTGAAGAAAGCCGGCAAGGTCAGTCGCGGCTGGCTGGGCGTGGTGA 7455

Qy     914 ttcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcat 973
           | ||||||||   | | | | | ||  |||| |||||  || || || ||   |||||
Db    7456 TCCAGGAAGTGAACAAGGATCTCGCCGAGTCCTTCGGCCTCGACAAGCCGTCCGGCGCGC 7515

Qy     974 tgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgaca 1033
           || | ||  || | |  . | | |||| || |     |  ||||||||| ||||||
Db    7516 TGGTGGCGCAACTGGTGGAAGACGGCCCGGCGGCCAAGGGTGGCCTGCAGGTGGGCGATG 7575

Qy    1034 tcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcg 1093
           |  ||||||||||  |||||      ||       |||| |||||| ||   | |||| |
Db    7576 TGATCCTCAGCCTGAACGGCCAGTCGATCAACGAGTCCGCCGACCTGCCGCACCTGGTGG 7635

Qy    1094 gcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcga 1145
           ||  ||| | |||||  |  |  ||| ||| | |||    ||||| ||| |
Db    7636 GCAACATGAAGCCGGGCGACAAGATCAACCTGGACGTGATTCGCAACGGCCA 7687
```

```
RESULT  10
AVU30799
LOCUS       AVU30799     2027 bp    DNA             BCT       31-MAY-1996
DEFINITION  Azotobacter vinelandii MucC (mucC) and MucD (mucD) genes, complete
            cds.
ACCESSION   U30799
VERSION     U30799.1  GI:1345102
KEYWORDS    .
SOURCE      Azotobacter vinelandii.
  ORGANISM  Azotobacter vinelandii
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Azotobacter.
REFERENCE   1  (bases 1 to 2027)
  AUTHORS   Martinez-Salazar,J.M., Moreno,S., Najera,R., Boucher,J.C.,
            Espin,G., Soberon-Chavez,G. and Deretic,V.
```

```
  TITLE     Characterization of the genes coding for the putative sigma factor
            AlgU and its regulators MucA, MucB, MucC, and MucD in Azotobacter
            vinelandii and evaluation of their roles in alginate biosynthesis
  JOURNAL   J. Bacteriol. 178 (7), 1800-1808 (1996)
  MEDLINE   96178940
REFERENCE   2  (bases 1 to 2027)
  AUTHORS   Martinez-Salazar,J.M., Moreno,S., Najera,R., Boucher,C., Espin,G.,
            Soberon-Chavez,G. and Deretic,V.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUN-1995) Jaime M. Martinez-Salazar, Microbiology,
            UTHSCSA, 7703 Floyd Curl Drive, San Antonio, TX 78284, USA
FEATURES             Location/Qualifiers
     source          1. .2027
                     /organism="Azotobacter vinelandii"
                     /strain="UW 136"
                     /db_xref="taxon:354"
     gene            1. .471
                     /gene="mucC"
     CDS             1. .471
                     /gene="mucC"
                     /note="similar to Photobacterium sp. ORF4 product, encoded
                     by GenBank Accession Number L41667"
                     /codon_start=1
                     /transl_table=11
                     /product="MucC"
                     /protein_id="AAB01512.1"
                     /db_xref="GI:1345103"
                     /translation="MIEEPGRVVAIEREAVWIEALDEAACPGCTASVGCGRWSFDWLE
                     TPRRRRRIRALSVMPLRVGDAVVVGIREDFLLRSTLAIYLLPLLGLFVCTIAIQQLGF
                     GESLMIPAGFLGFVSVWAMVRRISRRQMNDPTLQPVVLRALPGASEGQVRQAAF"
     CDS             483. .1904
                     /note="similar to Brucella abortus HtrA protein encoded by
                     GenBank Accession Number U07352"
                     /codon_start=1
                     /transl_table=11
                     /function="serine protease"
                     /product="MucD"
                     /protein_id="AAB01513.1"
                     /db_xref="GI:1345104"
                     /translation="MSKVGFKSFASVLAGALLLGQSLFVQAQLPEFTSLVEEASPAVV
                     NISTRQKLPDRSTVQGLPDLEGLPPLFREFLERSIPQLPRTPDNGRQREAHSLGSGFI
                     ISPDGYVLTNNHVVADADEIIVRLSDRSELEAELVGADPLTRCSFVEGQGFESPHSQT
                     GTYRPIESRGMGSGHRFPFRFRSFRDCGHHQCHGAKPAERELVPFIQTDVAINPGNSG
                     GPLFDLDGRVIGINSQIFTRSGGFMGLSFAIPIEVAMGVADQLKATGKVARGWLGVII
                     QEVNKDLAESFGLDRPAGALVAQVLEDGPADKGGLQVGDVILSLDGHPIVMSADLPHL
                     VGGLKPGAAANLEVVRDGKRRNIAITVGALPEEGNGVQPSIAGTEQSSNRLGVTVTEL
                     TAEQKKSLDLKGGVVIREVLNGPAALIGLRPGDVVTHLNNQPIDSAKTFAEVAGALPK
                     GRSVSMRVLRQGRASFITFKLAE"
BASE COUNT      366 a    529 c    658 g    474 t
ORIGIN
```

```
  Query Match             19.7%;  Score 274.8;  DB 2;  Length 2027;
  Best Local Similarity   58.9%;  Pred. No. 2.9e-41;
  Matches  527;  Conservative    0;  Mismatches  362;  Indels    5;  Gaps    3;

Qy     295 ttctacgaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagca 354
           |||   || ||  |  | |||  | | ||| |  | ||     |  ||  |  | |
Db     693 TTCAGGGAGTTCCTGGAGCGCAGCATTCCGCAACTTCCGCGTACTCCGGATAACGGGCGG 752

Qy     355 gatgacggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctg 414
           |    |  |       | | || ||||| ||||||||      | | |||||  | ||
Db     753 CAGCGTGAGGCGCACTCCCTGGGCTCGGGTTTCATCATTTCTCCAGATGGCTATGTTCTA 812

Qy     415 accaatacccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgc 474
           |||||  | ||| || || |||| |   |     ||||  || |   |   |||      |
Db     813 ACCAACAACCATGTGGTGGCCGATGCCGATGAAATCATCGTGCGTTTGTCCGATCGCAGT 872

Qy     475 gaatataccgccaaactcatcggttcggatg-tccaatccgatgtcgccctttctgaaaat 533
           ||       ||| | ||  ||||  | |||  ||  | ||||||| ||  |  ||||  |
Db     873 GAGCTCGAGGCCGAGCTGGTCGGGGCCGATCCTCTTACCCGATGTAGCTTTGTTGAAGGT 932

Qy     534 cgacgcaacggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccggg 593
```

```
           | | |   || |    || |||   |||||| | ||   | || |  | |||| |   ||
Db     933 CAAGGGTTCGAAT---CTCCCCACAGTCAAACTGGGACGTACCGACCAATTGAGAGTCGG 989

Qy     594 cgaatgggtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcat 653
           ||||||||    |||||||| | || ||||| || || |   |||||| || |||||
Db     990 GGAATGGGTTCTGGCCATCGGTTCCCCTTTCGGTTTCGATCATTCCGTGACTGCGGGCAT 1049

Qy     654 cgtgtccgccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccga 713
           | |    ||||  ||  |||||||||| ||||| || ||    || |||||||| |||||
Db     1050 CATCAGTGCCACGGGGCGAAGCCTGCCGAACGAGAGTTA-GTTCCTTTCATCCAGACCGA 1108

Qy     714 cgttgccatcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgt 773
           || |||||||||||| || || |||||||| ||||| ||| |  |  | |||| |||| |
Db     1109 TGTGGCCATCAATCCCGGTAACTCCGGCGGGCCCGCTCTTCGATCTGGATGGACGGGTCAT 1168

Qy     774 cggcatcaattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccat 833
           |||||||| || || |||| || |||    || || |||||||||| | || || || ||
Db     1169 AGGCATCAACTCCCAGATATTCACCCGGTCGGGGGGCTTCATGGGCTTGTCTTTCGCGAT 1228

Qy     834 cccgattgacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacg 893
           || || || |||||||||   || ||||| ||| ||||   ||| ||||| ||    ||
Db     1229 TCCCATCGAGGTTGCCATGGGCGTGGCCGATCAGTTGAAGGCCACTGGCAAGGTTGCTCG 1288

Qy     894 cggacaactgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtct 953
           |||     |||| || || ||||||||||   | | | | ||||  |||| || |||||
Db     1289 CGGTTGGTTGGGAGTAATCATTCAGGAAGTCAACAAGGATCTGGCTGAGTCCTTTGGTCT 1348

Qy     954 ggataaagccagcggcgcattgattgccaaaatcctttccggcagccccgcagaacgtgc 1013
           ||||    |   ||| || ||| | ||| |  || |     |   | || || ||    |
Db     1349 GGATCGGCCGGCCGGAGCGTTGGTCGCCCAGGTCTTGCAGGATGGACCGGCGGACAAGGG 1408

Qy     1014 cggcctgcaggcgggcgacatcgtcctcagcctcgacggcggagaaatacgtttcttccgg 1073
           |||  |||||  |||||  || | |||||||| || ||       ||       || |
Db     1409 CGGTTTGCAGGTCGGCGATGTCATTCTCAGCCTAGATGGTCATCCTATTGTGATGTCGGC 1468

Qy     1074 cgaccttcccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatg 1133
           ||| || ||     |||| || |   | |  || ||    | || ||  |||| ||
Db     1469 CGATCTGCCGCATCTGGTTGGGGGGCTCAAACCCGGGGCTGCAGCCAATCTCGAGGTGGT 1528

Qy     1134 gcgcaaaggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgcgagca 1187
           |||  | |||  |    |  |  |||     ||  | ||   |    | |||
Db     1529 GCGTGACGGCAAGCGGAGGAACATCGCTATCACTGTCGGGGCCTTGCCGGAGGA 1582
```

```
RESULT  11
AE004037
LOCUS       AE004037    12245 bp    DNA             BCT       06-JUL-2000
DEFINITION  Xylella fastidiosa, section 183 of 229 of the complete genome.
ACCESSION   AE004037 AE003849
VERSION     AE004037.1  GI:9107394
KEYWORDS    .
SOURCE      Xylella fastidiosa.
  ORGANISM  Xylella fastidiosa
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xylella.
REFERENCE   1  (bases 1 to 12245)
  AUTHORS   Silvestri,M.L., Siqueira,W.J., de Souza,A.A., de Souza,A.P.,
            Terenzi,M.F., Truffi,D., Tsai,S.M., Tsuhako,M.H., Vallada,H., Van
            Sluys,M.A., Verjovski-Almeida,S., Vettore,A.L., Zago,M.A., Zatz,M.,
            Meidanis,J. and Setubai,J.C.
  TITLE     The genome sequence of the plant pathogen Xylella fastidiosa. The
            Xylella fastidiosa Consortium of the Organization for Nucleotide
            Sequencing and Analysis, Sao Paulo, Brazil
  JOURNAL   Nature 406 (6792), 151-157 (2000)
  MEDLINE   20365717
REFERENCE   2  (bases 1 to 12245)
  AUTHORS   Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,
            Alvarenga,R., Alves,L.M.C., Araya,J.E., Baia,G.S., Baptista,C.S.,
            Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.S.,
            Bueno,M.R.P., Camargo,A.A., Camargo,L.E.A., Carraro,D.M.,
            Carrer,H., Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C.R.,
            Costa-Neto,C.M., Coutinho,L.L., Cristofani,M., Dias-Neto,E.,
            Docena,C., El-Dorry,H., Facincani,A.P., Ferreira,A.J.S.,
```

Ferreira,V.C.A., Ferro,J.A., Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Furlan,L.R., Garnier,M., Goldman,G.H., Goldman,M.H.S., Gomes,S.L., Gruber,A., Ho,P.L., Hoheisel,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P., Krieger,J.E., Kuramae,E.E., Laigret,F., Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A., Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N., Madeira,H.M.F., Marino,C.L., Marques,M.V., Martins,E.A.L., Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C., Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A., Nascimento,A.L.T.O., Netto,L.E.S., Nhani Jr.,A., Nobrega,F.G., Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C., Palmieri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V., de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V., Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M., Tsuhako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S., Vettore,A.L., Zago,M.A., Zatz,M., Meidanis,J. and Setubal,J.C.

TITLE Direct Submission

JOURNAL Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP 13083-970, Brazil

REMARK

FEATURES Location/Qualifiers

source 1. .12245
/organism="Xylella fastidiosa"
/db_xref="taxon:2371"

gene 80. .1624
/gene="XF2241"

CDS 80. .1624
/gene="XF2241"
/note="similar to GI|1184684 (percent identity: 47 %/query alignment coverage: 91.2 %/subject alignment coverage: 98.9 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -48"
/codon_start=1
/transl_table=11
/product="periplasmic protease"
/protein_id="AAF85040.1"
/db_xref="GI:9107395"
/translation="MNSRIHTRCFGLLAITLPLAACGAQHNNSPTPTAPSTPIVTPST
TPAPQLVAGLPDFTQLVDQVGPGVVNIETVITRKKVGKRRGIPLDNDIPEFFRRFFGP
DFQMPNQPRGGQDDEGGIAGRGMGSGFIISKDGYILTNHHVITGASEVTIKLTDRREF
KAKIIGSDEQYDVALLKIDAKNLPTVRIGDSSSLKSGQWVVAIGSPFGLDHSVTAGIV
SALGRSTSDDQRYVPFIQTDVPINQGNSGGPLLNTRGEVIGINSQIFSASGGYMGISF
AIPINLAINAAEQIRKTGKVQRSMLGVEIGPIDALKAQGLGLPDSRGALVNNIPPHSP
AAKAGIEVGDVIRSVNGKVISSFSDLPPLIGMMPPGTKATLGIIRDGKPREIVVTLSA
LNQDATASEDEDDTAPSKPETSANVELLGLQVENLSAAERERLASSQNAKGGVRITAV
TAPWARNANPPLVEGLVILRIGRTPVSNVAELNRVLGNYKKGDVIMLLVTNGRGTFYS
ALKAGG"

gene 1662. .1826
/gene="XF2242"

CDS 1662. .1826
/gene="XF2242"
/note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF85041.1"
/db_xref="GI:9107396"
/translation="MRLGVPSFLVGNGTPRPNIINTQVFLSTLTLNHAIILRYPDDGS
GAPSPCHPIR"

gene 1808. .3616
/gene="XF2243"

CDS 1808. .3616
/gene="XF2243"
/note="similar to SP|P07682 (percent identity: 69 %/query alignment coverage: 99.5 %/subject alignment coverage: 100.0 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="GTP binding protein"
/protein_id="AAF85042.1"
/db_xref="GI:9107397"
/translation="MSSDPMRNIRNFSIIAHVDHGKSTLADRIIQLCGGLEAREMEAQ
VLDSNPIERERGITIKAQSVSLLYKAQDGQNYHLNLIDTPGHVDFSYEVSRSLAACEG
ALLVVDASQGVEAQSVANCYTAVEQGLEVVPILNKIDLPTADTERAKAEIEAVIGIDA
SEAVAVSAKTGLYVEQVLEAIVQRIPAPQPRDTEKLQALIIDSWFDNYLGVVSLVRVM
QGEITPGAKLLVMSTGRSHPVDAVGVFTPKRKTLAKLTAGEVGWVTASIKDVHGAPVG
DTLTLANDPAPKPLPGFQEVQPRVFAGLFPVDAEDYPDLREALEKLRLNDAALRFEPE
NSEAMGFGFRCGFLGMLHMEIVQERLEREYDLNLITTAPTVIYEVLKTDGTLVAMDNP
AKMPPINQINEIREPIIRSNILTPPEYVGAVITLCEEKRGSQIGITYLGNQVQVAYEL
PMAEVVLDFFDKLKSVTRGYASLDYHFLRFQEGPFVRVDTLINGDRVDALSVIVHRHQ
AERRGRELCEKMKDLIPRQMFDVAIQAAIGSQIISRSTVKAMRKNVLAKCYGGDISRK
KKLLEKQKEGKKRMKQIGRVEIPQEAFLAVLQIDNK"

gene 3717. .4517
/gene="XF2244"

CDS 3717. .4517
/gene="XF2244"
/note="similar to SP|P00803 (percent identity: 40 %/query alignment coverage: 113.5 %/subject alignment coverage: 93.2 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="signal peptidase I"
/protein_id="AAF85043.1"
/db_xref="GI:9107398"
/translation="MRMKWFEIILVVLTLGTGIISLLDKLFLAKRRAARAGLLDTEPP
IIDYCRSFFPVLAAVLILRSFIAEPYKIPSSSMMPNLLIGDFILVNKFAYGLRLPITN
TKIIPIGKPNRGDVAVFKPPHKPDENWIKRVIGLPGDRIGFHGDTLYINGEPVKYILK
GQYTGKNAGVPDPTLLIEELPNHPHTILESIGRGREEGEGEWVVPPGQYFVMGDNRDN
SEDSRFWIKTHFLPEQNLRGKAFLIWLNCENWLCKGSFDPSRIGTVIR"

gene 4514. .4927
/gene="XF2245"

CDS 4514. .4927
/gene="XF2245"
/note="hypothetical protein; identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF85044.1"
/db_xref="GI:9107399"
/translation="MNERLVITKENLVKRTQYGITLRSTLIFTIIIFFTVHFAIHAFL
VYQKYYSLQTAMKSLKKELGNSDLDPSTIEELFLKRINIDSFDNVKKENIKFEKIRDG
WKMKINFEILRKRIGNMDLVGTFDTEQDLTSRDIK"

gene 4959. .5597
/gene="XF2246"

CDS 4959. .5597
/gene="XF2246"
/note="similar to SP|P05797 (percent identity: 50 %/query alignment coverage: 98.1 %/subject alignment coverage: 92.0 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder/Start codon shift: -93"
/codon_start=1
/transl_table=11
/product="ribonuclease III"
/protein_id="AAF85045.1"
/db_xref="GI:9107400"
/translation="MFTDPSLLLQALRHCSAGTPHNERLEFLGDSVVNLLIAEALFQR
WPRADEGALTRARSELVRETSLASIARTMQLGEQLILGPGELKSGGHRRDSILADAVE
AVIAAIYLDADLATCRTVVLPWFETALTALPVGKPEKDPKTRLQEWLQARQWSLPVYE
LIFESGDPHTKHFRVSCTLGELKLRTEGEGSSRRLAEQDAASHAIDQLDSNK"

gene 5594. .6490
/gene="XF2247"

CDS 5594. .6490
/gene="XF2247"
/note="similar to SP|P43728 (percent identity: 51 %/query alignment coverage: 97.3 %/subject alignment coverage: 96.0 %); identified by sequence similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1

```
                     /transl_table=11
                     /product="GTP binding protein"
                     /protein_id="AAF85046.1"
                     /db_xref="GI:9107401"
                     /translation="MTQTVPYRCGRIAVIGRPNVGKSTLTNALVGTKISIVSNRPQTT
                     RHRLLGIATFPEGQIVLVDTPGLHREQKHPMNRLMNRTARGSLEDVDAALLVTESTHW
                     NEEDTLAYNLLNDTGIPVVLVINKIDRFKDKSALLPFLTHINENHTFTTIHPVSALKR
                     KGLETLVSDLLALLPEGDPMFSEDEITDRSQRFLASELVREQVMRQLGEELPYATTVE
                     IEYFTENTGLFRIGALIWVERESQKAIVIGKGGARLKEIGVKARQQMERLFQTKVFLE
                     TWVRVRKDWSNNEAALKTFGYE"
     gene            6661. .7392
                     /gene="XF2248"
     CDS             6661. .7392
                     /gene="XF2248"
                     /note="similar to SP|P15027 (percent identity: 35 %/query
                     alignment coverage:  97.1 %/subject alignment coverage:
                     97.5 %); identified by sequence similarity; putative; ORF
                     located using Glimmer/RBSfinder"
                     /codon_start=1
                     /transl_table=11
                     /product="DNA repair protein RecO"
                     /protein_id="AAF85047.1"
                     /db_xref="GI:9107402"
                     /translation="MLIEHEVAFVLHVRPWRETSLLVEVLTQAYGRLGLIARGVQGLK
                     KQTLRAALQPLQWIRFSAIQRGELGQLRQAEALDTAPRLKGEAMLASFYINELLLRLV
                     PRHAPVNELYLAYSQTRERLRTSDSLAWSLRLFEREILETLGVGFNLECDANGTPLDP
                     ATRYVLDPLEGPRRLLSEHNNAKRRDTATGHVLLALARNQIPNTNDLAGLRRSMRAVL

  Query Match             18.7%;  Score 260.4;  DB 1;  Length 12245;
  Best Local Similarity   56.3%;  Pred. No. 8.9e-39;
  Matches  573;  Conservative    0;  Mismatches  431;  Indels   14;  Gaps    4;

Qy    159 gctgcccgactttgcccaactggttcaaagcgaaggcccggcagtcgtcaatatt----c 214
          | |||| || ||  |||||||||  |    | ||| || |  || ||||| |||    |
Db    238 GTTGCCAGATTTCACCCAACTGGTCGACCAAGTAGGGCCTGGCGTGGTCAACATTGAAAC 297

Qy    215 aggcagcccccgccccgcgcacccaaaacggcagcggcaatgccgaaaccgattccgacc 274
          | ||   ||||       |   ||||| |   |   |   |   ||        |||
Db    298 CGTCATTACCCGTAAGAAAGTCGGGAAACGTCGCGGCATACCGCTTGACAACGATATACC 357

Qy    275 cgcttgccgacagcgacccgttctacgaatttttcaaacgcctcgtcccgaacatgcccg 334
          ||  |     | |||   | ||   | |  ||| || | |||    |     | || | |
Db    358 CGAGTTTTTCCGGCGTTTCTTTGGCCCAGATTTCCAGATGCCGAACCAACCTCGTGGCGG 417

Qy    335 aaatcccccaagaagaagcagatgacggcggattgaacttcggttcgggcttcatcatca 394
              ||   | || || | || |  ||  ||   |    | || || ||||||||||
Db    418 ----CCAGGATGACGAGGGAGGTATCGCTGGTCGGGGAATGGGCTCTGGCTTCATCATTT 473

Qy    395 gcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatgggcagtatcaaag 454
           |||| |||||||||| || |||||    |||||| |  | ||   | ||    ||| |
Db    474 CCAAAGACGGCTACATACTCACCAACCATCACGTCATCACTGGCGCGAGCGAGGTCACAA 533

Qy    455 tcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatccg 514
          ||    | |  ||    |||||||  |  || ||| |||||||||   || |  || | ||
Db    534 TCAAATTGACTGATCGCCGCGAATTCAAAGCAAAAATCATCGGTAGCGACGAACAGTACG 593

Qy    515 atgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtcaaaatcggcaatc 574
          |||| ||  | || |||||||| ||||    || | ||||||  |||    |||||| |
Db    594 ATGTTGCTTTACTCAAAATCGATGCAA---AAAATCTACCCACCGTGCGTATCGGCGACT 650

Qy    575 ccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgcccttcggctttgaca 634
          |||       | ||| ||||  | |||||||  ||| ||||||  | ||||||| || ||
Db    651 CCAGTTCACTTAAATCGGGTCAGTGGGTCGTTGCGATCGGCTCCCCCTTCGGATTAGATC 710

Qy    635 acagcgtgaccgccggcatcgtgtccgcc---aaaggcagaagcctgcccaacgaaagct 691
          ||   || || || || || ||   ||||     | |||| | |     | |  || |||
Db    711 ACTCAGTCACTGCAGGGATAGTCAGCGCCCTGGGACGCAGCACCAGTGACGATCAACGCT 770

Qy    692 acacacccttcatccaaaccgacgttgccatcaatccgggcaattccggcggcccgctgt 751
          |    || || ||||||||||||||  |||| || | ||| ||||| || || ||   ||
Db    771 ATGTCCCATTTATCCAAACCGACGTACCCATTAACCAGGGAAATTCTGGTGGTCCATTGC 830

Qy    752 tcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatacagccgcagcggcggat 811
```

```
          | ||      |||  | ||  | || ||||| ||||||||||    |  |   ||||| |
Db    831 TGAATACCCGAGGTGAAGTGATTGGTATCAACTCGCAAATATTTTCCGCCTCAGGCGGCT 890

Qy    812 tcatgggcatctcctttgccatcccgattgacgttgccatgaatgtcgccgaacagctga 871
          ||||||  ||   |||||||||| || ||   |  | || || || |  |||||||  |
Db    891 ACATGGGTATTAGCTTTGCCATTCCAATCAATTTGGCAATCAACGCTGCCGAACAAATCC 950

Qy    872 aaaacaccggcaaagtccaacgcggacaactgggcgtgattattcaggaagtatcctacg 931
            || || |||||||| ||||||  |     || || |||   ||         |   |
Db    951 GTAAAACAGGCAAAGTGCAACGCAGTATGCTCGGTGTGGAAATCGGCCCTATCGACGCAC 1010

Qy    932 gtttggcacagtcgttcggtctggataaagccagcggcgcattgattgccaaaatccttc 991
              || ||     | ||  |     |   | | |||||| || |    || || | |
Db   1011 TCAAAGCTCAAGGCCTAGGCTTACCGGACAGCCGAGGCGCACTGGTCAATAACATTCCAC 1070

Qy    992 ccggcagccccgcagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgacg 1051
          ||  |||||| || |     || ||  |  | |  || ||| || ||| |    |  |||
Db   1071 CCCACAGCCCGGCCGCCAAAGCAGGGATTGAAGTCGGTGACGTCATCCGCTCAGTGAACG 1130

Qy   1052 gcggagaaatacgttcttccggcgaccttcccgtcatggtcggcgccattacgccgggaa 1111
          ||  || |||  ||   | | |||| || ||     || |||||   ||  |||| || |
Db   1131 GCAAAGTAATCAGTAGCTTCAGCGATCTGCCTCCACTGATCGGCATGATGCCGCCAGGCA 1190

Qy   1112 aagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagccaagct 1169
             ||| || ||| ||  |   ||  | ||| ||     |   |||   | ||  ||
Db   1191 CCAAAGCCACCCTGGGTATTATCCGTGACGGCAAACCGCGCGAGATCGTCGTCACTCT 1248
```

```
RESULT  12
AX011580
LOCUS       AX011580      390 bp    DNA             BCT       18-SEP-2000
DEFINITION  Sequence 9 from Patent WO9955872.
ACCESSION   AX011580
VERSION     AX011580.1  GI:9998109
KEYWORDS    .
SOURCE      Neisseria meningitidis.
  ORGANISM  Neisseria meningitidis
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Ruelle,J.L.
  TITLE     Basb013 dna and proteins from neisseria meningitidis
  JOURNAL   Patent: WO 9955872-A 04-NOV-1999;
            RUELLE JEAN LOUIS (BE); SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES             Location/Qualifiers
     source          1. .390
                     /organism="Neisseria meningitidis"
                     /db_xref="taxon:487"
BASE COUNT      106 a    120 c    102 g     62 t
ORIGIN


  Query Match             18.7%;  Score 260.2;  DB 2;  Length 390;
  Best Local Similarity   96.8%;  Pred. No. 2e-38;
  Matches  276;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy   1111 aaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagccaagctg 1170
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db      1 AAAGAAGTCAGCCTCGGCGTATGGCGCAAAGGCGAAGAAATCACAATCAAAGTCAAGCTG 60

Qy   1171 ggcaacgccgccgagcataccggcgcatcatccaaaacagatgaagccccctacaccgaa 1230
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGCAACGCCGCCGAGCATATCGGCGCATCATCCAAAACAGATGAAGCCCCCTACACCGAA 120

Qy   1231 cagcaatccggtacgttctcggtcgaatccgcaggcattacccttcagacacataccgac 1290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGCAATCCGGTACGTTCTCGGTCGAATCCGCAGGCATTACCCTTCAGACACATACCGAC 180

Qy   1291 agcagcggcaaacacctcgtcgtcgtacgggtttccgacgcggcagaacgcgcaggctta 1350
           |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGCAGCGGCGGACACCTCGTCGTCGTACGGGTTTCCGACGCGGCAGAACGCGCAGGCTTG 240

Qy   1351 aggcacggcgacgaaatcctagccgtcagggcaagtccccgtcaa 1395
```

```
          |||| |||||||||||| || |||||| ||||||||||||||||||
Db    241 AGGCGCGGCGACGAAATTCTTGCCGTC-GGGCAAGTCCCCGTCAA 284


RESULT  13
AF293977
LOCUS       AF293977     1362 bp    DNA             BCT       30-AUG-2000
DEFINITION  Aeromonas hydrophila htrA-like serine protease (prtS1) gene,
            complete cds.
ACCESSION   AF293977
VERSION     AF293977.1  GI:9945003
KEYWORDS    .
SOURCE      Aeromonas hydrophila.
  ORGANISM  Aeromonas hydrophila
            Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
            Aeromonas.
REFERENCE   1  (bases 1 to 1362)
  AUTHORS   Lin,T.-N., Lin,T.-J. and Liou,C.-M.
  TITLE     Aeromonas hydrophila strain CKH-29 prtS1 gene complete cds
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1362)
  AUTHORS   Lin,T.-N., Lin,T.-J. and Liou,C.-M.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-2000) Agricultural Chemistry, National Taiwan
            University, No.1 Roosevelt Road Section 4, Taipei 106, Taiwan
FEATURES             Location/Qualifiers
     source          1. .1362
                     /organism="Aeromonas hydrophila"
                     /strain="CKH-29"
                     /db_xref="taxon:644"
     gene            1. .1362
                     /gene="prtS1"
     CDS             1. .1362
                     /gene="prtS1"
                     /codon_start=1
                     /transl_table=11
                     /product="htrA-like serine protease"
                     /protein_id="AAG03073.1"
                     /db_xref="GI:9945004"
                     /translation="MRKPLSVLSVLALSVGIAMSAAPVQAALPSLLANNQEMPSLAPV
                     LEQVTPAVVNISVSGKKITRQRLPEQFRFFFGPNMPDEQVSEQPFQALGSGVIIDAKK
                     GYVITNAHVVHEADEIKVNLKDGREYAAKKIGEDKQSDIALLQIKAEDLVQIKFADSD
                     ELRVGDYALAIGNPFGLGQTVTSGIVSALGRSGLNIENLENFIQTDAAINSGNSGGAL
                     LNLRGELIGINTAILGPNGGNIGIGFAIPSNMVRDLSEQIVKYGEVRRGQLGIIGTEL
                     TSEVAKTFGYNKKDGAFVNQVMPDSAAAKAGIKPGDIIVSIDGKAIRSFGELRAKIAT
                     MGADKQVALGLIRDGKEQTVKVTLKKADDSEILASALHPALEGAKLSTTSEPVSGVAV
                     SEIDPRSPAAASGLQKGDVIIGVNRLRINSLEELSKALKNKPDVLALNIQRGDSSLYL
                     VIR"
BASE COUNT      289 a    448 c    394 g    231 t
ORIGIN


  Query Match             17.6%;  Score 246;  DB 2;  Length 1362;
  Best Local Similarity   58.3%;  Pred. No. 6.4e-36;
  Matches  470;  Conservative    0;  Mismatches  330;  Indels    6;  Gaps    2;

Qy    382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
          | | |||||  |  ||| || |||||| |  | |||||  ||||||| ||   ||    |
Db    286 GTCATCATCGACGCCAAGAAGGGCTACGTGATCACCAACGCCCACGTGGTGCACGAGGCG 345

Qy    442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
          | |   ||||| |||   || || ||    || |||||  |||||||    |||||
Db    346 GACGAGATCAAGGTCAACCTGAAGGATGGCCGTGAATACGCCGCCAAGAAGATCGGCGAA 405

Qy    502 gatgtccaatccgatgtcgccccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
          ||    || |||||  |||| || ||| ||||| | |    | || || ||         ||
Db    406 GACAAGCAGTCCGACATCGCTCTGCTGCAAATCAAGG---CCGAGGATCTGGTTCAAATC 462

Qy    562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
          ||| ||| | |  |  |  |  ||       || || |  |     |||||||||   ||
Db    463 AAATTCGCCGACTCTGACGAGCTGCGGGTCGGTGACTATGCGCTGGCCATCGGCAACCCG 522

Qy    622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgc---caaaggcagaagcctg 678
          |||||  |  |  |  | ||| ||| ||||||||||    ||   |    ||||  ||||
Db    523 TTCGGTCTCGGCCAGACCGTCACCTCCGGCATCGTCAGTGCACTCGGTCGCAGCGGCCTC 582

Qy    679 cccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaattcc 738
            || ||| | |     |  |||||||| |||||  |  |||||||  ||||||| ||
Db    583 AACATCGAGAACCTGGAAAACTTCATCCAGACCGATGCGGCCATCAACTCGGGCAACTCG 642

Qy    739 ggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatacagc 798
          || ||| | ||| ||||| |    ||  || |  ||||||||||  |    ||     ||
Db    643 GGTGGCGCCCTGCTCAACCTGCGCGGCGAGCTGATCGGCATCAACACCGCCATCCTGGGC 702

Qy    799 cgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaatgtc 858
          | || |||||   ||| ||||||  |||||||||| |||    || | |      |  |
Db    703 CCCAACGGCGGCAACATCGGCATCGGCTTTGCCATTCCGTCCAACATGGTGCGCGACCTG 762

Qy    859 gccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattattcag 918
          ||||||||  |    ||   |||  ||||||  ||||| || |||||| | ||
Db    763 TCCGAACAGATCGTGAAATACGGTGAAGTCCGTCGCGGCCAGCTGGGCATCATAGGGACC 822

Qy    919 gaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattgatt 978
          ||  |  |||  |   ||||  || | |||||      | ||      ||||||  ||  |
Db    823 GAGCTCACCTCTGAAGTGGCCAAGACCTTCGGCTACAACAAGAAGGACGGCGCCTTCGTC 882

Qy    979 gccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatcgtc 1038
            | |  |  | || | |    |||| |     ||||| |  || ||||||||||| ||
Db    883 AACCAGGTGATGCCTGACTCTGCCGCCGCCAAGGCCGGCATCAAGCCGGGCGACATCATC 942

Qy   1039 ctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggcgcc 1098
          ||||| ||||||||   |  || || || | |||||| || |  | || | | | ||
Db    943 GTCAGCATCGACGGCAAGGCCATTCGCTCCTTCGGCGAGCTGCGGGCCAAGATTGCCACC 1002

Qy   1099 attacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatc 1158
          ||     | |  ||  | ||    || ||| |    ||| | ||| | ||    ||  ||
Db   1003 ATGGGTGCCGACAAGCAGGTGGCACTGGGCCTGATCCGCGATGGCAAGGAGCAGACCGTC 1062

Qy   1159 aaagccaagctgggcaacgccgccga 1184
          || | ||  |||   || || |  ||
Db   1063 AAGGTCACCCTGAAGAAGGCGGATGA 1088


RESULT  14
AE004859
LOCUS       AE004859    11461 bp    DNA             BCT       30-AUG-2000
DEFINITION  Pseudomonas aeruginosa PA01, section 420 of 529 of the complete
            genome.
ACCESSION   AE004859 AE004091
VERSION     AE004859.1  GI:9950678
KEYWORDS    .
SOURCE      Pseudomonas aeruginosa.
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
            Pseudomonas.
REFERENCE   1  (bases 1 to 11461)
  AUTHORS   Stover,C.K., Pham,X.Q., Erwin,A.L., Mizoguchi,S.D., Warrener,P.,
            Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M.,
            Garber,R.L., Goltry,L., Tolentino,E., Westbrock-Wadman,S., Yuan,Y.,
            Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R.,
            Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T.
  TITLE     Complete genome sequence of Pseudomonas aeruginosa PA01, an
            opportunistic pathogen
  JOURNAL   Nature 406 (6799), 959-964 (2000)
  MEDLINE   20437337
REFERENCE   2  (bases 1 to 11461)
  AUTHORS   Stover,C.K., Pham,X.-Q.T., Erwin,A.L., Mizoguchi,S.D., Warrener,P.,
            Hickey,M.J., Brinkman,F.S.L., Hufnagle,W.O., Kowalik,D.J.,
            Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E.,
            Westbrock-Wadman,S., Yuan,Y., Brody,L.L., Coulter,S.N.,
            Folger,K.R., Kas,A., Larbig,K., Lim,R.M., Smith,K.A., Spencer,D.H.,
            Wong,G.K.-S., Wu,Z., Paulsen,I.T., Reizer,J., Saier,M.H.,
            Hancock,R.E.W., Lory,S. and Olson,M.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2000) Department of Medicine and Genetics,
            University of Washington Genome Center, University Of Washington,
```

```
          Box 352145, Seattle, WA 98195, USA
FEATURES             Location/Qualifiers
     source          1. .11461
                     /organism="Pseudomonas aeruginosa"
                     /strain="PAO1"
                     /db_xref="taxon:287"
     gene            203. .1306
                     /gene="mltB1"
                     /note="PA4444"
     CDS             203. .1306
                     /gene="mltB1"
                     /codon_start=1
                     /transl_table=11
                     /product="membrane-bound lytic transglycosylase"
                     /protein_id="AAG07832.1"
                     /db_xref="GI:9950679"
                     /translation="MRRTALALPLFLLVSACSSEPTPPPKPAAKPQARTVISPRPVRQ
                     SVQPILPLRGDYANNPAAQHFIDRMVSQHGFNRQQLHDLFAQTQRLDWVIRLMDRQAP
                     TYTPPSGPNGAWLRYRKKFVTPGNVQNGVLFWDQYETDLQRASRVYGVPPEIIVGIIG
                     VETRWGRVMGKTRIIDALSTLSFSYPRRAEFFSGELEQFLLQARKEGTDPLALRGSYA
                     GAMGYGQFMPSSFTKYAVDFDGDGHIDLWNPRDAIGSVANYFKQHGWVSGDRVAVPAS
                     GRAPSLEDGFKTLYPLDVLASAGLRPQGPLGGHRQASLLRLDMGRNYQYWYGLPNFYV
                     ITRYNHSTHYAMAVWELGKEVDRVRHRSVVRQD"
     gene            complement(1293. .2051)
                     /gene="PA4445"
     CDS             complement(1293. .2051)
                     /gene="PA4445"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAG07833.1"
                     /db_xref="GI:9950680"
                     /translation="MAIALSTLVEEADRYLDAARIQDYCPNGLQVEGRPQVRRIVSGV
                     TASQALLDAAVEADADVVLVHHGYFWKGENPCVVGMKQRRLKTLLNNDISLLAYHLPL
                     DLHPEVGNNVQLARQLGFEVEGPLEPGNPRSIVLLGSLAEPMQPADFARHVRDALGRE
                     PLLVDAGQPIRRIAWCTGGAQGYIDQAIAAGVDAYLTGEVSEQTVHSARENGISFIAA
                     GHHATERYGVQALGDYLGKRFAIEHLFIDCPNPA"
     gene            2168. .3337
                     /gene="algW"
                     /note="PA4446"
     CDS             2168. .3337
                     /gene="algW"
                     /codon_start=1
                     /transl_table=11
                     /product="AlgW protein"
                     /protein_id="AAG07834.1"
                     /db_xref="GI:9950681"
                     /translation="MPKALRFLGWPVLVGVLLALLIIQHNPELVGLPRQEVHVEQAPL
                     LSRLQEGPVSYANAVSRAAPAVANLYTTKMVSKPSHPLFDDPMFRRFFGDNLPQQKRM
                     ESSLGSAVIMSAEGYLLTNNHVTAGADQIIVALRDGRETIAQLVGSDPETDLAVLKID
                     LKNLPAMTLGRSDG1RTGDVCLAIGNPFGVGQTVTMGIISATGRNQLGLNTYEDFIQT
                     DAAINPGNSGGALVDAAGNLIGINTAIFSKSGGSQGIGFAIPTKLALEVMQSIIEHGQ
                     VIRGWLGVEVKALTPELAESLGLGETAGIVVAGVYRDGPAARGGLLPGDVILTIDKQE
                     ASDGRRSMNQVARTRPGQKISIVVLRNGQKVNLTAEVGLRPPPAPAPQQKQDGGE"
     gene            complement(3380. .4435)
                     /gene="hisC1"
                     /note="PA4447"
     CDS             complement(3380. .4435)
                     /gene="hisC1"
                     /codon_start=1
                     /transl_table=11
                     /product="histidinol-phosphate aminotransferase"
                     /protein_id="AAG07835.1"
                     /db_xref="GI:9950682"
                     /translation="MSKFWSPFVKDLVPYVPGEQPKLSRLVKLNTNENPYGPSPQALA
                     AMQAELNDDLRLYPDPNGERLKQAVAAHYGVQANQVFVGNGSDEVLAHIFHGLFQHDL
                     PLLFPDVTYSFYPVYCGLYGIAHEKIALDERFRIRVEDYARPNGGIIFPNPNAPTGCL
                     LPLDAIEAMLKASPDSVVVVDEAYVDFGGESAIALVDRYPNLLVTQTLSKSRSLAGLR
                     VGLAVGHADLVEALERIKNSFNSYPLDRLAIAGAAAAFEDDAYFRRTCQAVIDSREAL
                     SASLQALGFEVLPSAANFVFARHPRHDAGQIASTLREQGVIVRHFKQARIDQFLRITI
                     GSPEQNQALLDALHFLK"
     gene            complement(4438. .5760)
                     /gene="hisD"
                     /note="PA4448"
     CDS             complement(4438. .5760)
                     /gene="hisD"
                     /codon_start=1
                     /transl_table=11
                     /product="histidinol dehydrogenase"
                     /protein_id="AAG07836.1"
                     /db_xref="GI:9950683"
                     /translation="MTAPFAIRRLNAADPDFGRHLDHLLSWESVSDDSVNQRVLDIIA
                     AVRSRGDAAVVEFTQRFDGLQAASMADLILPRERLELALTRITVAQREALEVAAERVR
                     SYHEKQKQGSWRYTEADGTVLGQQVTPLDRAGLYVPGGKASYPSSVLMNAIPAKVAGV
                     SEVVMVVPTPRGEINEIVLAAACIAGVDRVFTIGGAQAVAALAYGTESVPRVDKIVGP
                     GNIYVATAKRHVFGQVGIDMIAGPSEILVVCDGQTDPDWIAMDLFSQAEHDEDAQSIL
                     VSPDAAFLDRVADSIARLLPTMERAEIIRTSLEGRGALIQVADQAQACAVANRIAPEH
                     LELSVADPESWLPEIRHAGAIFMGRYTAEALGDYCAGPNHVLPTSGTARFSSPLGVYD
                     FQKRSSIINCSAEGASVLGRTASVLARGESLTAHARSAEYRILDEKEA"
     gene            complement(5886. .6521)
                     /gene="hisG"
                     /note="PA4449"
     CDS             complement(5886. .6521)
                     /gene="hisG"
                     /codon_start=1
                     /transl_table=11
                     /product="ATP-phosphoribosyltransferase"
                     /protein_id="AAG07837.1"
                     /db_xref="GI:9950684"
                     /translation="MLTIALSKGRILDDTLPLLAAAGIVPSENPDKSRKLIIPTSLSD
                     VRLLIVRATDVPTYVEHGAADLGVAGKDVLMEYGGQGLYEPLDLRIANCKLMTAGAIG
                     APEPKGRLRVATKFVNVAKRYYAEQGRQVDVIKLYGSMELAPLVGLADKIIDVVDTGN
                     TLRANGLEPQELIATISSRLVVNKASMKMQHGRIQSLIDTLRDAVEARHRH"
     gene            complement(6539. .7804)
                     /gene="murA"
                     /note="PA4450"
     CDS             complement(6539. .7804)
                     /gene="murA"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-N-acetylglucosamine
                     1-carboxyvinyltransferase"
                     /protein_id="AAG07838.1"
                     /db_xref="GI:9950685"
                     /translation="MDKLIITGGNRLDGEIRISGAKNSALPILAATLLADTPVTVCNL
                     PHLHDITTMIELFGRMGVQPIIDEKLNVEVDASSIKTLVAPYELVKTMRASILVLGPM
                     LARFGEAEVALPGGCAIGSRPVDLHIRGLEAMGAQIEVEGGYIKAKAPAGGLRGGHFF
                     FDTVSVTGTENLMMAAALANGRTVLQNAAREPEVVDLANCLNAMGANVQGAGSDTIVI
                     EGVKRLGGARYDVLPDRIETGTYLVAAAATGGRVKLKDTDPTILEAVLQKLEEAGAHI
                     STGSNWIELDMKGNRPKAVNVRTAPYPAFPTDMQAQFISMNAVAEGTGAVIETVFENR
                     FMHVYEMNRMGAQILVEGNTAIVTGVPKLKGAPVMATDLRASASLVIAGLVAEGDTLI
                     DRIYHIDRGYECIEEKLQLLGAKIRRVPG"
     gene            complement(7825. .8064)
                     /gene="PA4451"
     CDS             complement(7825. .8064)
                     /gene="PA4451"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAG07839.1"
                     /db_xref="GI:9950686"
                     /translation="MQAEDVKSLLQEKLPGTQVEVEGEGCNFQLNLISDELAGQSPVK
                     RQQQVYAHLNAWIADGSIHAVTMKFFSRAAWAERS"
     gene            complement(8181. .8489)
                     /gene="PA4452"
     CDS             complement(8181. .8489)
                     /gene="PA4452"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAG07840.1"
                     /db_xref="GI:9950687"
                     /translation="MSQASLREGAAGELQLAGVLDYSSGPALREQGGRLIRASQAAEL
                     VVDCSAVERSSSVGISLLLAFIRDARKAGKVLSVRALPDDMREIAKVSSLLEILPLQE
                     "
     gene            complement(8486. .9133)
```

```
                     /gene="PA4453"
     CDS             complement(8486. .9133)
                     /gene="PA4453"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAG07841.1"
                     /db_xref="GI:9950688"
                     /translation="MLTLLRRGLLVFLAAFPLLSMAAPTPQQVVQGTVDELLSDIKAN
                     KAAYKADPQKLYATLDRILGPVVDAEGIAKSVMTVKYSRQASPEQIKRFEEVFKNSLM
                     QFYGNALLEYDNQDIRVLPSSAKPSDDRASVNMEIRDSKGTVYPVSYTMTNLAGGWKV
                     RNVIINGINIGKLFRDQFADTMQKNRNDLEKTIAGWGEVVAKAKETAKAEEAGAK"
     gene            complement(9145. .9618)
                     /gene="PA4454"
     CDS             complement(9145. .9618)
                     /gene="PA4454"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAG07842.1"
                     /db_xref="GI:9950689"
                     /translation="MQTRTLEIGVGLFLLAGLLALLLLALRVSGLSVGNAGDTYKVYA
                     YFDNIAGVTVRGKVTLAGVTIGKVTAVDLDRDSYTGRVTMEINQNVNNLPVDSTASIL

  Query Match             15.6%;  Score 217.4;  DB 1;  Length 11461;
  Best Local Similarity   55.9%;  Pred. No. 7.7e-31;
  Matches  475;  Conservative    0;  Mismatches  366;  Indels    8;  Gaps    3;

Qy     367 ttgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccac 426
           | || | |||| ||||   | ||||| |||    | |||||| | |||||||| | ||||
Db    2477 TCGAGCCTCGGCTCGGCGGTGATCATGAGCGCGGAAGGCTACCTGCTGACCAACAACCAC 2536

Qy     427 gtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgcc 486
           ||    || ||    | |   |||| ||     |   ||||  ||||||   | ||||
Db    2537 GTGACCGCTGGCGCCGACCAGATCATCGTGGCCTTGCGCGACGGCCGCGAAACCATCGCC 2596

Qy     487 aaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaa 546
           | | ||||    ||     || |||| | ||| | ||||| ||||||   |   |
Db    2597 CAGTTGGTCGGCAGCGACCCGGAAACCGACCTGGCCGTGCTGAAGATCGACCTTA---AG 2653

Qy     547 gagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgct 606
           | || || |   | |   |||||    || |     |     | |||||      |
Db    2654 AACCTGCCGGCGATGACCCTCGGCCGCTCCGACGGCATTCGCACCGGCGACGTCTGCCTC 2713

Qy     607 gccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaa 666
           |||||||||   || ||||| | | | | | |||||||   |||||| |   ||||
Db    2714 GCCATCGGCAACCCGTTCGGCGTCGGCCAGACCGTGACCATGGGCATCATCAGCGCCACC 2773

Qy     667 ggcagaagcctgcccaac---gaaagctacacaccccttcatccaaaccgacgttgccatc 723
           ||  | | || || |   |    | | ||||  |  |||||||| |||||||  || |||
Db    2774 GGACGCAACCAGCTCGGCCTGAACACCTACGAAGACTTCATCCAGACCGACGCGGCGATC 2833

Qy     724 aatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaat 783
           || || ||||| |||||||||| |||||  |  ||       ||  |  | ||||||||||
Db    2834 AACCCCGGCAACTCCGGCGGCGCGCTGGTGGACGCTGCCGGCAACCTGATCGGCATCAAC 2893

Qy     784 tcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgattgac 843
           ||   || | |   |     |||||| | |  ||| |||  ||| ||||||||||   |
Db    2894 ACGGCGATCTTCTCCAAGTCCGGCGGCTCCCAGGGTATCGGCTTCGCCATCCCGACCAAG 2953

Qy     844 gttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactg 903
           | ||| || | |||    | | |   |  |||| | ||    |||||     ||
Db    2954 CTGGCCCTGGAGGTCATGCAGTCGATCATCGAGCACGGCCAGGTGATCCGCGGCTGGCTC 3013

Qy     904 ggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaagcc 963
           |||||     |  |||   |  ||   |  ||||  ||||| | || || |   || ||
Db    3014 GGCGTCGAGGTCAAGGCGCTGACCCCGGAACTGGCGGAGTCGCTGGGCCTCGGCGAAACC 3073

Qy     964 agcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcag 1023
           |||     |  | |||    ||  || || | | || || | ||| | ||||||| |
Db    3074 GCCGGGATCGTCGTCGCCGGCGTCTATCGCGACGGTCCGGCGGCACGCGGCGGCCTGCTG 3133

Qy    1024 gcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttccc 1083
           |||||||  |  |||| | | |||||      |||    |      | ||  |
Db    3134 CCGGGCGATGTGATCCTGACCATCGACAAGCAGGAAGCCAGCGACGGCCGCCGCTCGATG 3193

Qy    1084 gtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggc 1143
           |  ||| |    ||     |||||| |  |  ||||| ||| ||  |||||| ||
Db    3194 AACCAGGTGGCGCGCACCCGTCCGGGACAGAAGATCAGCATCGTGGTGCTGCGCAACGGA 3253

Qy    1144 gaagaaatcacaatcaaagccaagctgggc--aacgccgccgagcataccggcgcatcat 1201
           |  |  |||   | |  ||| || | |||      ||||||     |||||| |  ||
Db    3254 CAGAAGGTCAACCTGACCGCCGAGGTCGGCCTGCGTCCGCCGCCGGCACCGGCTCCACAG 3313

Qy    1202 ccaaaacag 1210
           |  |||||| 
Db    3314 CAGAAACAG 3322


RESULT  15
RMU31512
LOCUS       RMU31512     1974 bp    DNA             BCT       05-MAR-1996
DEFINITION  Rhizobium meliloti RmDEGP (degP) gene, complete cds.
ACCESSION   U31512
VERSION     U31512.1  GI:951169
KEYWORDS    .
SOURCE      Rhizobium meliloti.
  ORGANISM  Sinorhizobium meliloti
            Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
            Rhizobiaceae; Sinorhizobium.
REFERENCE   1  (bases 1 to 1974)
  AUTHORS   Glazebrook,J., Ichige,A. and Walker,G.C.
  TITLE     Genetic analysis of Rhizobium meliloti bacA-phoA fusion results in
            identification of degP: two loci required for symbiosis are closely
            linked to degP
  JOURNAL   J. Bacteriol. 178 (3), 745-752 (1996)
  MEDLINE   96146524
REFERENCE   2  (bases 1 to 1974)
  AUTHORS   Glazebrook,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-1995) Jane Glazebrook, Department of Molecular
            Biology, Massachusetts General Hospital, Wellman 10, MGH, Boston,
            MA 02114, USA
FEATURES             Location/Qualifiers
     source          1. .1974
                     /organism="Sinorhizobium meliloti"
                     /strain="Rm1021"
                     /db_xref="taxon:382"
     gene            236. .1690
                     /gene="degP"
     CDS             236. .1690
                     /gene="degP"
                     /codon_start=1
                     /transl_table=11
                     /product="RmDEGP"
                     /protein_id="AAC43669.1"
                     /db_xref="GI:951170"
                     /translation="MSTIKAPRPSLKTVLKTTTVAGLAAVPVTTGLPAEVAQSFAEAV
                     RVQAPAVAVSPMWSTPFRRRSSPSACRHVNASATMKATSPSISAAAGSRTCRKTIRCG
                     VSSANSLRVKMTVPIVGATAAVRVAKVVSVRGRKAPASSSPKTVTSSPTTTSSPTART
                     FTVIMNDGTELDAKLVGKDSRTDLAVLKVDDKRKFTYVSFADDEKVRVGDWVVAVGNP
                     FGLGGTVTAGIISARGRDIGSGPYDDYLQVDAAVNRGNSGGPTFNLSGEVVGINTAIF
                     SPSGGNVGIAFAIPASVAKDVVDSLIKDGTVSRGWLGVQIQPVTKDIAESLGLSEANG
                     ALVVEPQAGSPGEKAGIKNGDVVTALNGEPVKDPRDLARRVAALRPGSTAEVTLWRSG
                     KSETVNLEIGTLPSDAKEPAPATGEAQPDEGQAGEEALADLGLTVTPSEDGKGVTIAS
                     VDPDSDAGDRGLKEGEKIVSVNNQEVNRQTTFSR"
BASE COUNT      378 a    626 c    626 g    344 t
ORIGIN


  Query Match             15.2%;  Score 212.4;  DB 2;  Length 1974;
  Best Local Similarity   53.4%;  Pred. No. 9.4e-30;
  Matches  539;  Conservative    0;  Mismatches  461;  Indels   10;  Gaps    4;

Qy     376 ggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgcc 435
           || || |||||| |||||| | || |||| ||| || | ||||| | |||||||||  ||
```

```
Db    643 GGCTCCGGCTTCTTCATCACCGAAGACGGTTACCTCGTCACCAACAACCACGTCGTCTCC 702

Qy    436 ggtatgg-gcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcat 494
          |       | |   |||  ||  |  | ||||||     |||      ||||||| ||| |
Db    703 GACGGCTCGGACCTTCACCGTTATCATGAACGACGGAACCGAGCTCGACGCCAAGCTCGT 762

Qy    495 cggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacc 554
          |||   |||   ||   |||| | ||| | || ||  |||| |  | |    || |  |
Db    763 CGGCAAGGACAGCCGGACCGACCTTGCCGTGCTCAAGGTCGATGACAAGCGCAAGTTCAC 822

Qy    555 cgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcgg 614
          |   ||||   |||   |   | | ||  ||      ||||| ||||| |  ||  ||||
Db    823 CTATGTCAGCTTCGCGGACGACGAGAAGGTGCGCGTCGGCGACTGGGTGGTCGCTGTCGG 882

Qy    615 cgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcagaag 674
          |   |||||||| | | |  ||  || || |||||||| | |||||    ||| |
Db    883 CAATCCCTTCGGCCTCGGCGGCACGGTCACTGCCGGCATCATCTCCGCTCGGGGCCGCGA 942

Qy    675 cctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaa 734
          |   |  ||   |||     |   | ||   |||||  ||  | || |  |||||
Db    943 TATCGGCTCCGGTCCCTATGACGATTATCTGCAGGTCGACGCAGCGGTGAACCGCGGCAA 1002

Qy    735 ttccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatata 794
          |||||| || |||   |||||  |    ||  |||| ||||| |||||  ||   || |
Db   1003 TTCCGGTGGTCCGACCTTCAATCTTTCCGGCGAGGTGGTCGGTATCAACACGGCCATCTT 1062

Qy    795 cagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaa 854
          | |  ||||||     | || ||| |||||||| |||||       || || | | |
Db   1063 TTCGCCCTCCGGCGGCAATGTCGGTATCGCCTTTGCAATCCCCGCCTCGGTCGCGAAGGA 1122

Qy    855 tgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattat 914
          |||||  ||    || |  ||   |||||  ||    |||||    || || ||    ||
Db   1123 TGTCGTGGACTCCCTCATCAAGGACGGCACGGTGTCGCGCGGCTGGCTTGGTGTCCAGAT 1182

Qy    915 tcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcatt 974
          |||   ||  |  | | | | ||  ||||| |||| ||     |||| | ||||||  |
Db   1183 CCAGCCGGTGACGAAGGATATCGCCGAGTCGCTCGGCCTTTCCGAAGCGAACGGCGCCCT 1242

Qy    975 gattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacat 1034
          | | | |    |  | |||   || |  |||   |||||  |  ||   |||||| |
Db   1243 CGTCGTCGAGCCTCAGGCGGGCTCGCCGGGTGAAAAGGCCGGGATCAAGAACGGCGACGT 1302

Qy   1035 cgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcgg 1094
          |||       ||| ||||||      |     | | | |||||  |       || |
Db   1303 CGTGACGGCACTCAACGGCGAGCCGGTCAAGGATCCGCGGGACCTCGCACGTCGCGTTGC 1362

Qy   1095 cgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcac 1154
          ||| |    || ||       |      |   | | |||||||  ||| ||      ||
Db   1363 GGCCCTGCGCCCCGGCTCCACTGCAGAGGTTACCCTGTGGCGCAGCGGCAAATCCGAAAC 1422

Qy   1155 aatcaaagccaagctgggca---acgccgccgagcataccggcgcatcatccaaaacaga 1211
            ||||   | || | ||||      |   |||    |  |   |  || |    || |
Db   1423 GGTCAACCTCGAGATCGGCACGCTGCCGAGCGATGCCAAGGAACCCGCACCGGCGACCGG 1482

Qy   1212 tgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatccgcaggcattac 1271
          |||||| |    |  ||||  |||  | ||   |    || | |    |   ||| | ||
Db   1483 TGAAGCGCAGCCCGACGAAGGGCAGGCAGGCGAGGAGGCGCTGGCCGACCTCGGCCTGAC 1542

Qy   1272 cct---tcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgggtttccga 1328
            |    | | |  |   || ||   ||| ||    |  |  |||||  | || ||| ||
Db   1543 GGTGACCCCGTCGGAGGACGGCAAGGGCGTCACGATCGCCTCCGTCGATCCGGATTCGGA 1602

Qy   1329 cgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtca 1378
          ||| |  || |||   ||| | | | | |||||  | ||| |  ||||||
Db   1603 CGCCGGCGATCGC---GGCCTGAAGGAAGGCGAGAAGATCGTCTCCGTCA 1649


Search completed: March 22, 2001, 04:06:30
Job time: 4575 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 03:20:15 ; Search time 91.55 Seconds
(without alignments)
5724.188 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaaataccaata.....tcagggcaagtccccgtcaa 1395

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*
1: /cgn2_2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /cgn2_2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /cgn2_2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /cgn2_2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /cgn2_2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /cgn2_2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /cgn2_2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /cgn2_2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /cgn2_2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /cgn2_2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /cgn2_2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /cgn2_2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /cgn2_2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /cgn2_2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /cgn2_2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /cgn2_2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /cgn2_2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /cgn2_2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /cgn2_2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /cgn2_2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /cgn2_2/gcgdata/geneseq/geneseqn/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1395	100.0	1395	21	Z93414	NGSP polypeptide c
2	1393.4	99.9	1395	21	Z51538	Neisseria meningit
3	1373.4	98.5	1500	21	Z54510	Neisseria gonorrhe
4	1324.4	94.9	1326	21	Z51539	Neisseria meningit
5	1320.6	94.7	1500	21	Z33307	Neisseria meningit
6	1315.8	94.3	1500	21	Z54511	Neisseria meningit
7	1313.8	94.2	1500	21	Z54512	Neisseria meningit
8	1263	90.5	1500	21	Z33306	Neisseria meningit
9	1258.2	90.2	1500	21	Z33305	Neisseria meningit
10	1242	89.0	1242	21	Z93415	NGSP polypeptide c
11	1176.2	84.3	1347	21	Z51533	Neisseria meningit
12	1061	76.1	1110	21	Z33308	Neisseria meningit
13	364.8	26.2	1436	20	X98241	Nucleotide sequenc
14	260.2	18.7	390	21	Z33309	Neisseria meningit
15	176.4	12.6	1854	19	V29528	Escherichia coli s
16	175.6	12.6	1980	12	Q14416	S. typhimurium htr
17	154.2	11.1	1791	17	T04403	Heat shock protein
18	153	11.0	153	21	Z93416	NGSP coding sequen
c 19	150.4	10.8	1038602	20	Z01425	Complete genome se
20	139.4	10.0	2894	17	T10423	H. influenzae SB33
21	138.6	9.9	153	21	Z51542	Neisseria meningit
c 22	137.8	9.9	1830121	17	T42063	Haemophilus influe
23	128.8	9.2	1529	19	X14164	H. pylori GHPO 536
24	124	8.9	1540	20	V90845	Nucleotide sequenc
25	120.8	8.7	1484	20	V90921	Nucleotide sequenc
c 26	110.8	7.9	558	19	V31239	E. coli J96 pathog
27	105	7.5	1185	21	Z29195	cDNA encoding C-te
28	105	7.5	1443	21	Z29176	Recombinant cDNA e
29	105	7.5	1705	20	Z41293	Human normal ovari
30	105	7.5	2036	19	V29540	Homo sapiens D8725
31	105	7.5	2075	18	T75444	cDNA encoding oste
32	105	7.5	2205	19	V04680	Human presenilin i
33	85	6.1	607	20	X29130	Polynucleotide RTP
34	81.6	5.8	2040	21	Z52362	NSEQ gene-14 assoc
35	80.6	5.8	1302	21	Z48700	BASB011 coding seq
36	80.6	5.8	1302	21	Z48701	BASB011 coding seq
37	80.6	5.8	1302	21	Z48702	BASB011 coding seq
c 38	71	5.1	284	20	V90593	Nucleotide sequenc
39	68	4.9	1553	19	V43618	Human secreted pro
40	62.8	4.5	273	20	V90690	Nucleotide sequenc
41	60.6	4.3	1194	21	Z91868	Streptococcus pneu
42	60.6	4.3	15747	20	X13018	Enterococcus faeca
43	60.6	4.3	21338	19	V52153	Streptococcus pneu
44	60.4	4.3	9848	20	X20503	Polynucleotide seq
45	57.6	4.1	1835	19	V29524	Homo sapiens PSP1

ALIGNMENTS

RESULT 1
Z93414
ID Z93414 standard; DNA; 1395 BP.
XX
AC Z93414;
XX
DT 24-JUL-2000 (first entry)
XX
DE NGSP polypeptide coding sequence.
XX
KW NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW cellular matrix; adherence domain; ligand; detection; diagnosis;
KW screening; probe; primer; prophylaxis; therapy; ss.
XX
OS Neisseria gonorrhoeae.
XX
FH Key Location/Qualifiers
FT CDS 1..1395
FT /*tag= a
FT /product= NGSP polypeptide
XX
PN WO200012133-A1.
XX
PD 09-MAR-2000.
XX
PF 01-SEP-1999; 99WO-US20070.
XX
PR 01-SEP-1998; 98US-0098685.
XX
PA (ANTE-) ANTEX BIOLOGICS INC.
XX
PI Jackson WJ, Harris AM;
XX
DR WPI; 2000-237782/20.
DR P-PSDB; Y83150.

XX
PT Non-cytosolic NGSP polypeptide and polynucleotide sequence from
PT Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT infections
XX
PS Claim 38; Page 60-61; 68pp; English.
XX
CC The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and
CC Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC domains for extracellular matrix proteins. Using the NGSP polypeptide
CC as a vaccine produces antibodies which inhibit binding of N.
CC gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC subsequent invasion. The NGSP polypeptide and its peptide fragments
CC can be used to immunise an animal and produce an immune response.
CC They can also be used as ligands to detect antibodies elicited in
CC response to Neisseria infections and also as antigens or immunogens
CC for inducing Neisseria-specific antibodies which are useful in
CC immunoassays to detect Neisseria in biological specimens. Nucleotides
CC encoding NGSP or its fragments can be used as probes to identify
CC Neisseria in biological specimens by hybridization or polymerase
CC chain reaction amplification. The NGSP polypeptide can also be used
CC in screening assays to identify agents and compounds which useful as
CC diagnostic, prophylactic or therapeutic agents against Neisseria
CC infection.
XX
SQ Sequence 1395 BP; 359 A; 438 C; 353 G; 245 T; 0 other;

Query Match 100.0%; Score 1395; DB 21; Length 1395;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60

Qy      61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120

Qy     121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180

Qy     181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240

Qy     241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300

Qy     301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360

Qy     361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420

Qy     421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480

Qy     481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540

Qy     541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600

Qy     601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660

Qy     661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720

Qy     721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780

Qy     781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840

Qy     841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900

Qy     901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960

Qy     961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020

Qy    1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080

Qy    1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140

Qy    1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200

Qy    1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260

Qy    1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320

Qy    1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380

Qy    1381 gcaagtccccgtcaa 1395
          |||||||||||||||
Db    1381 gcaagtccccgtcaa 1395
```

RESULT 2
Z51538
ID Z51538 standard; DNA; 1395 BP.
XX
AC Z51538;
XX
DT 03-JUL-2000 (first entry)
XX
DE Neisseria meningitidis NMASP protein-2 encoding DNA.
XX
KW NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW Neisserial infection; meningitidis; septicaemia; ds.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers

```
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "NMASP protein"
FT                   /note= "Does not include stop codon"
FT                   /transl_except= (pos:22..24, aa:Leu)
FT                   /transl_except= (pos:49..51, aa:Ser)
FT                   /transl_except= (pos:64..66, aa:Asp)
FT                   /transl_except= (pos:124..126, aa:Lys)
FT                   /transl_except= (pos:172..174, aa:Val)
FT                   /transl_except= (pos:250..252, aa:Ser)
FT                   /transl_except= (pos:400..402, aa:Asp)
FT                   /transl_except= (pos:433..435, aa:Thr)
FT                   /transl_except= (pos:580..582, aa:Asp)
FT                   /transl_except= (pos:649..654, aa:Gly)
FT                   /transl_except= (pos:964..966, aa:Gly)
FT                   /transl_except= (pos:1021..1023, aa:Arg)
FT                   /transl_except= (pos:1162..1164, aa:Val)
FT                   /transl_except= (pos:1189..1191, aa:Ile)
FT                   /transl_except= (pos:1300..1302, aa:Gly)
FT                   /transl_except= (pos:1354..1356, aa:Arg)
FT                   /transl_except= (pos:1378..1395, aa: gqvpvndeagfrkamdkagk
FT                   nvpllimrrgntlfialnlq)
FT                   /partial
XX
PN   WO200012535-A2.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US19663.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ,  Harris AM;
XX
DR   WPI; 2000-256581/22.
DR   P-PSDB; Y70413.
XX
PT   Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT   antibodies, useful in vaccines against infection  -
XX
PS   Claim 41; Page 69; 75pp; English.
XX
CC   The present sequence is a DNA encoding Neisseria meningitidis NMASP
CC   protein. NMASP is a non-cytosolic protein, with antibacterial and
CC   antiinflammatory activity. It shows sequence similarity to E. coli
CC   DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC   antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC   anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
CC   may be used for diagnosis, therapy or prophylaxis of Neisserial
CC   infections such as, bacterial meningitidis and septicaemia.
XX
SQ   Sequence 1395 BP; 358 A; 439 C; 353 G; 245 T; 0 other;
```

```
  Query Match             99.9%;  Score 1393.4;  DB 21;  Length 1395;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1394;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tgcgaaaaggccggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420

Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720

Qy    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960

Qy    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020

Qy   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080

Qy   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140

Qy   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200

Qy   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260

Qy   1261 gcaggcattaccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380

Qy   1381 gcaagtccccgtcaa 1395
          |||||||||||||||
Db   1381 gcaagtccccgtcaa 1395


RESULT   3
Z54510
ID   Z54510 standard; DNA; 1500 BP.
XX
AC   Z54510;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria gonorrheae ORF 986 partial DNA sequence SEQ ID NO:2967.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria gonorrheae.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; Y75748.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 1389-1390; 1453pp; English.
XX
CC   Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC   novel Neisseria meningitis and N. gonorrheae polynucleotides and
CC   polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
CC   primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1500 BP; 384 A; 472 C; 379 G; 265 T; 0 other;
```

```
  Query Match             98.5%;  Score 1373.4;  DB 21;  Length 1500;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1388;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||
Db      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctcccaggc 60

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420

Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720

Qy    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960

Qy    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
```

```
Qy   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080

Qy   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140

Qy   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1141 ggcgaagaaatcacaatcaaagccaagctgggccacgccgccgagcataccggcgcatca 1200

Qy   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260

Qy   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          |||||||||||||||||||||||||||||||| |||||||||||||| || ||| ||
Db   1321 gtttccgacgcggcagaacgcgcaggcttaaggcgcggcgacgaaatcctcgcggtc-gg 1379

Qy   1381 gcaagtccccgtcaa 1395
          |||||||||||||||
Db   1380 gcaagtccccgtcaa 1394
```

```
RESULT   4
Z51539
ID   Z51539 standard; DNA; 1326 BP.
XX
AC   Z51539;
XX
DT   03-JUL-2000  (first entry)
XX
DE   Neisseria meningitidis NMASP protein-3 encoding DNA.
XX
KW   NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW   anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW   Neisserial infection; meningitidis; septicaemia; ds.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1326
FT                   /*tag=  a
FT                   /product= "NMASP protein"
FT                   /note= "Does not include stop codon"
FT                   /transl_except= (pos:55..57, aa:Lys)
FT                   /transl_except= (pos:103..105, aa:Val)
FT                   /transl_except= (pos:181..183, aa:Ser)
FT                   /transl_except= (pos:331..333, aa:Asp)
FT                   /transl_except= (pos:364..366, aa:Thr)
FT                   /transl_except= (pos:511..513, aa:Asp)
FT                   /transl_except= (pos:580..585, aa:Gly)
FT                   /transl_except= (pos:895..897, aa:Gly)
FT                   /transl_except= (pos:952..954, aa:Arg)
FT                   /transl_except= (pos:1093..1095, aa:Val)
FT                   /transl_except= (pos:1120..1122, aa:Ile)
FT                   /transl_except= (pos:1231..1233, aa:Gly)
FT                   /transl_except= (pos:1285..1287, aa:Arg)
FT                   /transl_except= (pos:1309..1311, aa:Gly)
FT                   /transl_except= (pos:1312..1314, aa:Gln)
FT                   /transl_except= (pos:1315..1317, aa:Val)
FT                   /transl_except= (pos:1321..1326, aa: vndeagfrkamdkagknvpl
FT                   limrrgntlfialnlq)
FT                   /partial
XX
PN   WO200012535-A2.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US19663.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ,  Harris AM;
XX
DR   WPI; 2000-256581/22.
DR   P-PSDB; Y70414.
XX
PT   Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT   antibodies, useful in vaccines against infection  -
XX
PS   Claim 41; Page 73; 75pp; English.
XX
CC   The present sequence is a DNA encoding Neisseria meningitidis NMASP
CC   protein. NMASP is a non-cytosolic protein, with antibacterial and
CC   antiinflammatory activity. It shows sequence similarity to E. coli
CC   DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC   antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC   anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
CC   may be used for diagnosis, therapy or prophylaxis of Neisserial
CC   infections such as, bacterial meningitidis and septicaemia.
XX
SQ   Sequence 1326 BP; 342 A; 422 C; 334 G; 228 T; 0 other;
```

```
  Query Match             94.9%;  Score 1324.4;  DB 21;  Length 1326;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1325;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     70 gcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgcatcgaacac 129
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gccggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgcatcgaacac 60

Qy    130 accaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactggttcaaagc 189
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 accaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactggttcaaagc 120

Qy    190 gaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaaaacggcagc 249
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaaaacggcagc 180

Qy    250 ggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctacgaatttttc 309
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctacgaatttttc 240

Qy    310 aaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgacggcggattg 369
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 aaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgacggcggattg 300

Qy    370 aacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtc 429
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 aacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtc 360

Qy    430 gttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaa 489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 gttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaa 420

Qy    490 ctcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagag 549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ctcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagag 480

Qy    550 ctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgcc 609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgcc 540

Qy    610 atcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggc 669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 atcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggc 600
```

```
Qy    670 agaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccg 729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 agaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccg 660

Qy    730 ggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaa 789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaa 720

Qy    790 atatacagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgcc 849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 atatacagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgcc 780

Qy    850 atgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtg 909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 atgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtg 840

Qy    910 attattcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggc 969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 attattcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggc 900

Qy    970 gcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggc 1029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 gcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggc 960

Qy   1030 gacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatg 1089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatg 1020

Qy   1090 gtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaa 1149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaa 1080

Qy   1150 atcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatcatccaaaaca 1209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 atcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatcatccaaaaca 1140

Qy   1210 gatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatccgcaggcatt 1269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 gatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatccgcaggcatt 1200

Qy   1270 acccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgggtttccgac 1329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 acccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgggtttccgac 1260

Qy   1330 gcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagggcaagtccc 1389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagggcaagtccc 1320

Qy   1390 cgtcaa 1395
          ||||||
Db   1321 cgtcaa 1326
```

```
RESULT   5
Z33307
ID   Z33307 standard; DNA; 1500 BP.
XX
AC   Z33307;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Neisseria meningitidis strain H44/76 BASB013 nucleotide sequence.
XX
KW   Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW   antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW   invasive bacterial disease; antibacterial; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9955872-A1.
XX
PD   04-NOV-1999.
XX
PF   20-APR-1999;   99WO-EP02765.
XX
PR   23-APR-1998;   98GB-0008734.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-052809/04.
DR   P-PSDB; Y52995.
XX
PT   Novel polynucleotides and polypeptides from Neisseria meningitis used
PT   to prepare vaccines against bacterial infections -
XX
PS   Claim 12; Page 78-79; 94pp; English.
XX
CC   The present sequence encodes a BASB013 polypeptide isolated from
CC   Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC   employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence 1500 BP; 385 A; 471 C; 382 G; 262 T; 0 other;
```

```
  Query Match             94.7%;  Score 1320.6;  DB 21;  Length 1500;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 1355;  Conservative    0;  Mismatches   39;  Indels    1;  Gaps    1;

Qy      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          |||||||||||||||||||| |||||||||| |||||||||||||||| ||||||||||
Db      1 gtgttcaaaaaataccaatacctcgctttggcagcactgtgtgccgcctcgctggcaggc 60

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtaqaacgc 120
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tgcgacaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtaqaacgc 120

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    121 atcgaacacaccaaagacgacggcagcgtcagtatgctgctgcccgactttgcccaactg 180

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
          |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gttcaaagtgaaggtccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          |||||||||||||||||||||| |||||||||||| ||||||||| ||||||||||||||
Db    241 aacggcagcggcaatgccgaaaacgattccgacccgattgccgacaacgacccgttctac 300

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    301 gaatttttcaaacgcctcgtcccgaatatgcccgaaatcccccaagaagaagcagatgac 360

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    361 ggcggattgaacttcggttcgggcttcatcatcagcaaagacggctacatcctgaccaat 420
```

```
Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          |||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db    421 acccacgtcgttaccggcatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          |||||||||| |||||||||||||||||||||||||| ||||||||||||||||||||
Db    541 acggaagagctgcccgtcgtcaaaatcggcaatcccaaagatttgaaaccgggcgaatgg 600

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    601 gtcgccgccatcggcgcgcccttcggcttcgacaacagcgtgaccgccggcatcgtgtcc 660

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacaccttcatccaaaccgacgttgcc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gccaaaggcagaagcctgcccaacgaaagctacacaccttcatccaaaccgacgttgcc 720

Qy    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    721 atcaatccgggcaactccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
          || |||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db    781 aactcgcaaatatacagccgcagcggcggattcatgggcatttccttcgccatcccgatt 840

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          |||||||||||||||| |||||||||||||||||||| |||||||  |||| |||
Db    901 ctgggcgtgattattcaagaagtatcctacggtttggcacaatcgttcggcttggacaaa 960

Qy    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
          ||| |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db    961 gccggcggcgcactgattgccaaaatcctgcccggcagccccgcagaacgtgccggcctg 1020

Qy   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080

Qy   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 cccgttatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140

Qy   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
          |||||||||||||||||||| |||||||||||||||||||| ||||||||||
Db   1141 ggcgaagaaatcacaatcaaagtcaagctgggcaacgccgccgagcatatcggcgcatca 1200

Qy   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260

Qy   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db   1261 gcaggcattacccttcagacacataccgacagcagcggcggacacctcgtcgtcgtacgg 1320

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          |||||||||||||||||||||||||||||| |||| |||||||||||| || |||||| ||
Db   1321 gtttccgacgcggcagaacgcgcaggcttgaggcgcggcgacgaaattcttgccgtc-gg 1379

Qy   1381 gcaagtccccgtcaa 1395
          |||||||||||||||
Db   1380 gcaagtccccgtcaa 1394
```

```
RESULT   6
Z54511
ID   Z54511 standard; DNA; 1500 BP.
XX
AC   Z54511;
XX
DT   21-MAR-2000  (first entry)
XX
DE   Neisseria meningitidis ORF 986 partial DNA sequence SEQ ID NO:2969.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
XX
PD   11-NOV-1999.
XX
PF   30-APR-1999;   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C,  Galeotti C,  Grandi G,  Hickey E,  Masignani V,  Mora M;
PI   Petersen J,  Pizza M,  Rappuoli R,  Ratti G,  Scalato E,  Scarselli M;
PI   Tettelin H,  Venter JC;
XX
DR   WPI; 2000-062150/05.
DR   P-PSDB; Y75749.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics  -
XX
PS   Claim 7; Page 1390; 1453pp; English.
XX
CC   Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC   novel Neisseria meningitis and N. gonorrheae polynucleotides and
CC   polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
CC   primers used in the exemplification of the present invention. The
CC   polypeptides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1500 BP; 386 A; 469 C; 382 G; 263 T; 0 other;
```

```
  Query Match           94.3%;  Score 1315.8;  DB 21;  Length 1500;
  Best Local Similarity  96.9%;  Pred. No. 0;
  Matches 1352;  Conservative    0;  Mismatches   42;  Indels    1;  Gaps    1;

Qy      1 gtgttcaaaaaatacccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          |||||||||||||||||||| |||||||||| |||||||||| |||| ||||||||||
Db      1 gtgttcaaaaaatacccaatacctcgctttggcagcactgtgtgcagcctcgctggcaggc 60

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaaagaagcatccttcgtagaacgc 120
          ||||| |||||||||||||| ||||  ||||||||||||||||||||||||||||||||
Db     61 tgcgacaaggcaggcagcttcttcgtggcggacaaaaaagaagcatccttcgtagaacgc 120

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
```

```
Db    121 atcgaacacaccaaagacgacggcagcgtcagtatgctgctgcccgactttgcccaactg 180

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
Db    181 gttcaaagtgaaggtccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
Db    241 aacggcagcggcaatgccgaaaacgattccgacccgattgccgacaacgacccgttctac 300

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
Db    301 gaatttttcaaacgcctcgtcccgaatatgcccgaaatcccccaagaagaagcagatgac 360

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
Db    361 ggcggattgaacttcggttcgggcttcatcatcagcaaagacggctacatcctgaccaat 420

Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
Db    421 acccacgtcgttaccggcatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
Db    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
Db    541 acggaagagctgcccgtcgtcaaaatcggcaatcccaaagatttgaaaccgggcgaatgg 600

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
Db    601 gtcgccgccatcggcgcgcccttcggcttcgacaacagcgtgaccgccggcatcgtgtcc 660

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
Db    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720

Qy    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
Db    721 atcaatccgggcaactccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
Db    781 aactcgcaaatatacagccgcagcggcggattcatgggcatttccttcgccatcccgatt 840

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
Db    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
Db    901 ctgggcgtgattattcaagaagtatcctacggtttggcacaatcgttcggtttggacaaa 960

Qy    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
Db    961 gccggcggcgcactgattgccaaaatcctgcccggcagccccgcagaacgtgccggcctg 1020

Qy   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
Db   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080

Qy   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
Db   1081 cccgttatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140

Qy   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
Db   1141 ggcgaagaaatcacaatcaaagtcaagctgggcaacgccgccgagcatatcggcgcatca 1200

Qy   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
Db   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260

Qy   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
Db   1261 gcaggcattacccttcagacacataccgacagcagcggcggacacctcgtcgtcgtacgg 1320

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
Db   1321 gtttccgacgcggcagaacgcgcaggcttgaggcgcggcgacgaaattcttgccgtc-gg 1379

Qy   1381 gcaagtccccgtcaa 1395
Db   1380 gcaagtccccgtcaa 1394
```

RESULT 7
Z54512
ID Z54512 standard; DNA; 1500 BP.
XX
AC Z54512;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 986 partial DNA sequence SEQ ID NO:2971.
XX
KW Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR P-PSDB; Y75750.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 7; Page 1392; 1453pp; English.
XX
CC Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC novel Neisseria meningitis and N. gonorrheae polynucleotides and
CC polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
CC primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention

```
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 1500 BP; 382 A; 469 C; 382 G; 265 T; 2 other;

  Query Match             94.2%;  Score 1313.8;  DB 21;  Length 1500;
  Best Local Similarity   96.8%;  Pred. No. 0;
  Matches 1350;  Conservative    0;  Mismatches   44;  Indels    1;  Gaps    1;

Qy       1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
           |||||||||||||||||| |||||||||| ||||||||||||||| ||||||||||
Db       1 gtgttcaaaaaataccaatacctcgctttggcagcactgtgtgccgcctcgctggcaggc 60

Qy      61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaaacgc 120
           ||||| || || |||||||||||||||||||||||||||||||||||| |||||||||
Db      61 tgcgacaaagccggcagctttttcggtgcggacaaaaaagaagcatcctttgtagaacgc 120

Qy     121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
           ||| |||||||||||||||||||| |||||||||||||||||||||| |||||||
Db     121 atcaaacacaccaaagacgacggcagcgtcagtatgctgctgcccgactttgtccaactg 180

Qy     181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240

Qy     241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 aacggcagcagcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300

Qy     301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360

Qy     361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
           || |||||||||||||||||||||||||||||||||||| ||||||| || |||||||||
Db     361 ggnggattgaacttcggttcgggcttcatcatcagcaaagacggctatattctgaccaat 420

Qy     421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
           || |||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db     421 acgcacgtcgttaccggcatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480

Qy     481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540

Qy     541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
           |||||||||||| |||||||||||||||||||||||||| |||||||||||||||||||
Db     541 acggaagagctgcccgtcgtcaaaatcggcaatcccaaagatttgaaaccgggcgaatgg 600

Qy     601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
           ||||| |||||||||||||||||||||| |||||||||||||||||||| ||||||||
Db     601 gtcgccgccatcggcgcgcccttcggcttcgacaacagcgtgaccgccggcntcgtgtcc 660

Qy     661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720

Qy     721 atcaatccgggcaatttccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     721 atcaatccgggcaactccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780

Qy     781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
           || |||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db     781 aactcgcaaatatacagccgcagcggcggattcatgggcatttccttcgccatcccgatt 840

Qy     841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900

Qy     901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
           |||||||||||||||| |||||||||||||||||||||| ||||||||| |||| |||
Db     901 ctgggcgtgattattcaagaagtatcctacggtttggcacaatcgttcggtttggacaaa 960

Qy     961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
           ||| |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db     961 gccggcggcgcactgattgccaaaatcctgcccggcagccccgcagaacgtgccggcctg 1020

Qy    1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 cgggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080

Qy    1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 cccgttatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140

Qy    1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
           |||||||||||||||||||| |||||||||||||||||||||| ||||||||||
Db    1141 ggcgaagaaatcacaatcaaagtcaagctgggcaacgccgccgagcatatcggcgcatca 1200

Qy    1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260

Qy    1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
           |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db    1261 gcaggcattacccttcagacacataccgacagcagcggcggacacctcgtcgtcgtacgg 1320

Qy    1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
           |||||||||||||||||||||||||||||| |||| |||||||||||| || |||||| ||
Db    1321 gtttccgacgcggcagaacgcgcaggcttgaggcgcggcgacgaaattcttgccgtc-gg 1379

Qy    1381 gcaagtccccgtcaa 1395
           |||||||!|||||||
Db    1380 gcaagtccccgtcaa 1394


RESULT   8
Z33306
ID   Z33306 standard; DNA; 1500 BP.
XX
AC   Z33306;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Neisseria meningitidis strain ATCC 13090 BASB013 nucleotide sequence.
XX
KW   Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW   antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW   invasive bacterial disease; antibacterial; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9955872-A1.
XX
PD   04-NOV-1999.
XX
PF   20-APR-1999;   99WO-EP02765.
XX
PR   23-APR-1998;   98GB-0008734.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-052809/04.
DR   P-PSDB; Y52994.
XX
PT   Novel polynucleotides and polypeptides from Neisseria meningitis used
PT   to prepare vaccines against bacterial infections  -
XX
PS   Claim 12; Page 76; 94pp; English.
XX
CC   The present sequence encodes a BASB013 polypeptide isolated from
CC   Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC   employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
```

```
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence 1500 BP; 370 A; 469 C; 392 G; 269 T; 0 other;

  Query Match             90.5%;  Score 1263;  DB 21;  Length 1500;
  Best Local Similarity   94.6%;  Pred. No. 0;
  Matches 1319;  Conservative    0;  Mismatches   75;  Indels    1;  Gaps    1;

Qy      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          |||||||||||||||||||| |||||||||| |||||||||||||||| |||||||||
Db      1 gtgttcaaaaaataccaatacctcgctttggcagcactgtgtgccgcctcgctggcaggc 60

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
          ||||| || || ||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tgcgacaaagccggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          |||||||||||||||||||||||||| |||||||||||||||||||||||| |||||||
Db    121 atcgaacacaccaaagacgacggcagcgtcagtatgctgctgcccgactttgtccaactg 180

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    301 gaatttttcaaacgcctcgtcccgaatatgcccgaaatcccccaagaagaagcagatgac 360

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          |||||||||||||||||||||||||||||||||||||||| ||||||| || |||||||||
Db    361 ggcggattgaacttcggttcgggcttcatcatcagcaaagacggctatattctgaccaat 420

Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          || ||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db    421 acgcacgtcgttaccggcatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgccccttctgaaaatcgacgca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 accgccaaactcatcggttcggatgtccaatccgatgtcgccccttctgaaaatcgacgca 540

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db    541 acggaagagctgcccgtcgtcaaaatcggcaatcccaaagatttgaaaccgggcgaatgg 600

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||| |||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    601 gtcgccgccatcggcgcgcccttcggcttcgacaacagcgtgaccgccggcatcgtgtcc 660

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720

Qy    721 atcaatccgggcaatttccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          |||||||||||||| |||||||||||||||||||| | ||||||||||||||||||||
Db    721 atcaatccgggcaactccggcggcccgctgttcaacctgaaaggacaggtcgtcggcatc 780

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
          || |||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db    781 aactcgcaaatatacagccgcagcggcggattcatgggcatttccttcgccatcccgatt 840

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          |||||||||||||||| |||||||||||||||||||| |||||||| |||| |||
Db    901 ctgggcgtgattattcaagaagtatcctacggtttggcacaatcgttcggtttggacaaa 960

Qy    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
          ||| |||||||| ||||||||||||||| |||||||||||||||||||||||||||||||
Db    961 gccggcggcgcactgattgccaaaatcctgcccggcagccccgcagaacgtgccggcctg 1020

Qy   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080

Qy   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 cccgttatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140

Qy   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
          ||  | ||||||||  |    | ||| |||||||| ||  |||| || |||||  | ||
Db   1141 ggtaaggaaatcaccgttgccgtcaaactgggcaatgcttccgaacaaaccggttcctcg 1200

Qy   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
          ||| |  | ||  |||||||  ||  |||||||  ||||||||||||||||||||||||
Db   1201 tccgagccggacaaagccccttatgccgaacaccaatccggtacgttctcggtcgaatcc 1260

Qy   1261 gcaggcattacccttcagacacataccgacagcagcggcgaacacctcgtcgtcgtacgg 1320
          |||||||||||||||||||||||||||||||||||||||  ||  || |||||||| |||
Db   1261 gcaggcattacccttcagacacataccgacagcagcggcggacggcttgtcgtcgtgcgg 1320

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          ||||| |  |||||||||||||||||||| |||| |||||||||||||| |||||| ||
Db   1321 gtttcgggggcggcagaacgcgcaggcttgaggcgcggcgacgaaatccttgccgtc-gg 1379

Qy   1381 gcaagtccccgtcaa 1395
          |||||||||||||||
Db   1380 gcaagtccccgtcaa 1394


RESULT   9
Z33305
ID   Z33305 standard; DNA; 1500 BP.
XX
AC   Z33305;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Neisseria meningitidis strain ATCC 13090 BASB013 nucleotide sequence.
XX
KW   Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW   antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW   invasive bacterial disease; antibacterial; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9955872-A1.
XX
PD   04-NOV-1999.
XX
PF   20-APR-1999;   99WO-EP02765.
XX
PR   23-APR-1998;   98GB-0008734.
```

```
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-052809/04.
DR   P-PSDB; Y52993.
XX
PT   Novel polynucleotides and polypeptides from Neisseria meningitis used
PT   to prepare vaccines against bacterial infections -
XX
PS   Claim 15; Page 73-74; 94pp; English.
XX
CC   The present sequence encodes a BASB013 polypeptide isolated from
CC   Neisseria meningitidis. BASB013 polynucleotides and polypeptides may be
CC   employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence 1500 BP; 370 A; 468 C; 393 G; 269 T; 0 other;


  Query Match             90.2%;  Score 1258.2;  DB 21;  Length 1500;
  Best Local Similarity   94.3%;  Pred. No. 0;
  Matches 1316;  Conservative    0;  Mismatches   78;  Indels    1;  Gaps    1;

Qy      1 gtgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggc 60
          |||||||||||||||||||| |||||||||| ||||||||||||||| ||||||||||
Db      1 gtgttcaaaaaataccaatacctcgctttggcagcactgtgtgccgcctcgctggcaggc 60

Qy     61 tgcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120
          ||||| || || ||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tgcgacaaagccggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgc 120

Qy    121 atcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaactg 180
          |||||||||||||||||||||||| |||||||||||||||||||||||| |||||||
Db    121 atcgaacacaccaaagacgacggcagcgtcagtatgctgctgcccgactttgtccaactg 180

Qy    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaa 240

Qy    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 aacggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctac 300

Qy    301 gaatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgac 360
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    301 gaatttttcaaacgcctcgtcccgaatatgcccgaaatcccccaagaagaagcagatgac 360

Qy    361 ggcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaat 420
          |||||||||||||||||||||||||||||||||||||| ||||||| || |||||||||
Db    361 ggcggattgaacttcggttcgggcttcatcatcagcaaagacggctatattctgaccaat 420

Qy    421 acccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480
          || |||||||||| |||| ||||||||||||||||||||||||||||||||||||||||
Db    421 acgcacgtcgttaccggcatgggcagtatcaaagtcctgctcaacgacaagcgcgaatat 480

Qy    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 accgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgca 540

Qy    541 acggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgg 600
          |||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db    541 acggaagagctgcccgtcgtcaaaatcggcaatcccaaagatttgaaaccgggcgaatgg 600

Qy    601 gtcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtcc 660
          ||||| ||||||||   |||||||||||| ||||||||||||||||||||||||||||||
Db    601 gtcgccgccatcgggcggcccttcggcttcgacaacagcgtgaccgccggcatcgtgtcc 660

Qy    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 gccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcc 720

Qy    721 atcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatc 780
          |||||||||||||| |||||||||||||||||||| | ||||||||||||||||||||
Db    721 atcaatccgggcaactccggcggcccgctgttcaacctgaaaggacaggtcgtcggcatc 780

Qy    781 aattcgcaaatatacagccgcagcggcggattcatgggcatctcctttgccatcccgatt 840
          || ||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db    781 aactcgcaaatatacagccgcagcggcggattcatgggcatttccttcgccatcccgatt 840

Qy    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gacgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaa 900

Qy    901 ctgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaa 960
          |||||||||||||||| |||||||||||||||||||||||| ||||||||| |||| |||
Db    901 ctgggcgtgattattcaagaagtatcctacggtttggcacaatcgttcggtttggacaaa 960

Qy    961 gccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctg 1020
          ||| |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db    961 gccggcggcgcactgattgccaaaatcctgcccggcagccccgcagaacgtgccggcctg 1020

Qy   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 caggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgacctt 1080

Qy   1081 cccgtcatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 cccgttatggtcggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaa 1140

Qy   1141 ggcgaagaaatcacaatcaaagccaagctgggcaacgccgccgagcataccggcgcatca 1200
          ||  | ||||||||  |    | ||| |||||||| ||  |||| || |||||  | ||
Db   1141 ggtaaggaaatcaccgttgccgtcaaactgggcaatgcttccgaacaaaccggttcctcg 1200

Qy   1201 tccaaaacagatgaagccccctacaccgaacagcaatccggtacgttctcggtcgaatcc 1260
          ||| |  | ||  |||||| ||  |||||| ||||||||||||||||||||||||||||
Db   1201 tccgagccggacaaagccccttatgccgaacaccaatccggtacgttctcggtcgaatcc 1260

Qy   1261 gcaggcattacccttcagacacataccgacagcagcggcaaacacctcgtcgtcgtacgg 1320
          |||||||||||||||||||||||||||||||||||||||||  ||  || |||||||| |||
Db   1261 gcaggcattacccttcagacacataccgacagcagcggcggacggcttgtcgtcgtgcgg 1320

Qy   1321 gtttccgacgcggcagaacgcgcaggcttaaggcacggcgacgaaatcctagccgtcagg 1380
          ||||| |  |||||||||||||||||||| |||| |||||||||||||| |||||| ||
Db   1321 gtttcgggggcggcagaacgcgcaggcttgaggcgcggcgacgaaatccttgccgtc-gg 1379

Qy   1381 gcaagtccccgtcaa 1395
          |||||||||||||||
Db   1380 gcaagtccccgtcaa 1394


RESULT  10
Z93415
ID   Z93415 standard; DNA; 1242 BP.
XX
AC   Z93415;
XX
```

```
DT   24-JUL-2000  (first entry)
XX
DE   NGSP polypeptide coding sequence.
XX
KW   NGSP; polypeptide; peptide; vaccine; immune response; antibody;
KW   cellular matrix; adherence domain; ligand; detection; diagnosis;
KW   screening; probe; primer; prophylaxis; therapy; ss.
XX
OS   Neisseria gonorrhoeae.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1242
FT                   /*tag= a
FT                   /product= NGSP polypeptide
XX
PN   WO200012133-A1.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US20070.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ,  Harris AM;
XX
DR   WPI; 2000-237782/20.
DR   P-PSDB; Y83151.
XX
PT   Non-cytosolic NGSP polypeptide and polynucleotide sequence from
PT   Neisseria useful for diagnosis, prevention or treatment of Neisseria
PT   infections
XX
PS   Claim 38; Page 62-63; 68pp; English.
XX
CC   The NGSP polypeptide of N. gonorrhoeae has conserved Arg-Gly-Asp and
CC   Arg-Gly-Asn motifs near the C-terminus which function as adherence
CC   domains for extracellular matrix proteins. Using the NGSP polypeptide
CC   as a vaccine produces antibodies which inhibit binding of N.
CC   gonorrhoeae to the host's cellular matrix reducing attachment and/or
CC   subsequent invasion. The NGSP polypeptide and its peptide fragments
CC   can be used to immunise an animal and produce an immune response.
CC   They can also be used as ligands to detect antibodies elicited in
CC   response to Neisseria infections and also as antigens or immunogens
CC   for inducing Neisseria-specific antibodies which are useful in
CC   immunoassays to detect Neisseria in biological specimens. Nucleotides
CC   encoding NGSP or its fragments can be used as probes to identify
CC   Neisseria in biological specimens by hybridization or polymerase
CC   chain reaction amplification. The NGSP polypeptide can also be used
CC   in screening assays to identify agents and compounds which useful as
CC   diagnostic, prophylactic or therapeutic agents against Neisseria
CC   infection.
XX
SQ   Sequence 1242 BP; 316 A; 400 C; 312 G; 214 T; 0 other;
```

```
  Query Match             89.0%;  Score 1242;  DB 21;  Length 1242;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     154 atgctgctgcccgactttgcccaactggttcaaagcgaaggcccggcagtcgtcaatatt 213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atgctgctgcccgactttgcccaactggttcaaagcgaaggcccggcagtcgtcaatatt 60

Qy     214 caggcagcccccgccccgcgcacccaaaacggcagcggcaatgccgaaaccgattccgac 273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 caggcagcccccgccccgcgcacccaaaacggcagcggcaatgccgaaaccgattccgac 120

Qy     274 ccgcttgccgacagcgacccgttctacgaatttttcaaacgcctcgtcccgaacatgccc 333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ccgcttgccgacagcgacccgttctacgaatttttcaaacgcctcgtcccgaacatgccc 180

Qy     334 gaaatcccccaagaagaagcagatgacggcggattgaacttcggttcgggcttcatcatc 393
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gaaatcccccaagaagaagcagatgacggcggattgaacttcggttcgggcttcatcatc 240

Qy     394 agcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatgggcagtatcaaa 453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 agcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatgggcagtatcaaa 300

Qy     454 gtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatcc 513
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 gtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatcc 360

Qy     514 gatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtcaaaatcggcaat 573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 gatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtcaaaatcggcaat 420

Qy     574 cccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgcccttcggctttgac 633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgcccttcggctttgac 480

Qy     634 aacagcgtgaccgccggcatcgtgtccgccaaaggcagaagcctgcccaacgaaagctac 693
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 aacagcgtgaccgccggcatcgtgtccgccaaaggcagaagcctgcccaacgaaagctac 540

Qy     694 acacccttcatccaaaccgacgttgccatcaatccgggcaattccggcggcccgctgttc 753
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 acacccttcatccaaaccgacgttgccatcaatccgggcaattccggcggcccgctgttc 600

Qy     754 aacttaaaaggacaggtcgtcggcatcaattcgcaaatatacagccgcagcggcggattc 813
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aacttaaaaggacaggtcgtcggcatcaattcgcaaatatacagccgcagcggcggattc 660

Qy     814 atgggcatctcctttgccatcccgattgacgttgccatgaatgtcgccgaacagctgaaa 873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 atgggcatctcctttgccatcccgattgacgttgccatgaatgtcgccgaacagctgaaa 720

Qy     874 aacaccggcaaagtccaacgcggacaactgggcgtgattattcaggaagtatcctacggt 933
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 aacaccggcaaagtccaacgcggacaactgggcgtgattattcaggaagtatcctacggt 780

Qy     934 ttggcacagtcgttcggtctggataaagccagcggcgcattgattgccaaaatccttccc 993
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ttggcacagtcgttcggtctggataaagccagcggcgcattgattgccaaaatccttccc 840

Qy     994 ggcagccccgcagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgacggc 1053
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ggcagccccgcagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgacggc 900

Qy    1054 ggagaaatacgttcttccggcgaccttcccgtcatggtcggcgccattacgccgggaaaa 1113
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ggagaaatacgttcttccggcgaccttcccgtcatggtcggcgccattacgccgggaaaa 960

Qy    1114 gaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagccaagctgggc 1173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagccaagctgggc 1020

Qy    1174 aacgccgccgagcataccggcgcatcatccaaaacagatgaagccccctacaccgaacag 1233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 aacgccgccgagcataccggcgcatcatccaaaacagatgaagccccctacaccgaacag 1080

Qy    1234 caatccggtacgttctcggtcgaatccgcaggcattacccttcagacacataccgacagc 1293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 caatccggtacgttctcggtcgaatccgcaggcattacccttcagacacataccgacagc 1140

Qy    1294 agcggcaaacacctcgtcgtcgtacgggtttccgacgcggcagaacgcgcaggcttaagg 1353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 agcggcaaacacctcgtcgtcgtacgggtttccgacgcggcagaacgcgcaggcttaagg 1200

Qy    1354 cacggcgacgaaatcctagccgtcagggcaagtccccgtcaa 1395
           |||||||||||||||||||||||||||||||||||||||||||
Db    1201 cacggcgacgaaatcctagccgtcagggcaagtccccgtcaa 1242
```

```
RESULT  11
Z51533
ID   Z51533 standard; DNA; 1347 BP.
XX
AC   Z51533;
XX
DT   03-JUL-2000  (first entry)
XX
DE   Neisseria meningitidis NMASP protein-1 encoding DNA.
XX
KW   NMASP; non-cytosolic; antibacterial; antiinflammatory; cytotoxic;
KW   anti-NMASP antibody; vaccine; diagnosis; therapy; prophylaxis;
KW   Neisserial infection; meningitidis; septicaemia; ds.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1347
FT                   /*tag= a
FT                   /product= "Neisseria meningitidis protein"
FT                   /transl_except= (pos:499..501, aa: Xaa)
FT                   /note= "Xaa is unknown"
XX
PN   WO200012535-A2.
XX
PD   09-MAR-2000.
XX
PF   01-SEP-1999;   99WO-US19663.
XX
PR   01-SEP-1998;   98US-0098685.
XX
PA   (ANTE-) ANTEX BIOLOGICS INC.
XX
PI   Jackson WJ,  Harris AM;
XX
DR   WPI; 2000-256581/22.
DR   P-PSDB; Y70409.
XX
PT   Neisseria meningitidis NMASP polypeptide, nucleotide sequences and
PT   antibodies, useful in vaccines against infection  -
XX
PS   Claim 41; Page 63; 75pp; English.
XX
CC   The present sequence is a DNA encoding Neisseria meningitidis NMASP
CC   protein. NMASP is a non-cytosolic protein, with antibacterial and
CC   antiinflammatory activity. It shows sequence similarity to E. coli
CC   DegP (HtrA) protein. NMASP proteins can be used as ligands to detect
CC   antibodies elicited in response to N. meningitidis infections. Cytotoxic
CC   anti-NMASP antibodies can be used as vaccines. NMASP proteins and DNA
CC   may be used for diagnosis, therapy or prophylaxis of Neisserial
CC   infections such as, bacterial meningitidis and septicaemia.
CC   Note: The protein sequence represented in SEQ ID NO:2 of the
CC   specification is erroneous"
XX
SQ   Sequence 1347 BP; 338 A; 428 C; 344 G; 236 T; 1 other;

  Query Match             84.3%;  Score 1176.2;  DB 21;  Length 1347;
  Best Local Similarity   97.2%;  Pred. No. 0;
  Matches 1207;  Conservative    0;  Mismatches   34;  Indels    1;  Gaps    1;

Qy     154 atgctgctgcccgactttgcccaactggttcaaagcgaaggcccggcagtcgtcaatatt 213
           ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1 atgctgctgcccgactttgtccaactggttcaaagcgaaggcccggcagtcgtcaatatt 60

Qy     214 caggcagcccccgccccgcgcacccaaaacggcagcggcaatgccgaaaccgattccgac 273
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      61 caggcagcccccgccccgcgcacccaaaacggcagcagcaatgccgaaaccgattccgac 120

Qy     274 ccgcttgccgacagcgacccgttctacgaatttttcaaacgcctcgtcccgaacatgccc 333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ccgcttgccgacagcgacccgttctacgaatttttcaaacgcctcgtcccgaacatgccc 180

Qy     334 gaaatcccccaagaagaagcagatgacggcggattgaacttcggttcgggcttcatcatc 393
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gaaatcccccaagaagaagcagatgacggcggattgaacttcggttcgggcttcatcatc 240

Qy     394 agcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatgggcagtatcaaa 453
           |||||| ||||||| || |||||||||| ||||||||| |||| ||||||||||||||||
Db     241 agcaaagacggctatattctgaccaatacgcacgtcgttaccggcatgggcagtatcaaa 300

Qy     454 gtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatcc 513
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 gtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatcc 360

Qy     514 gatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtcaaaatcggcaat 573
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     361 gatgtcgcccttctgaaaatcgacgcaacggaagagctgcccgtcgtcaaaatcggcaat 420

Qy     574 cccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgcccttcggctttgac 633
           |||||| |||||||||||||||||||||||| |||||||||||||||||||||| |||
Db     421 cccaaagatttgaaaccgggcgaatgggtcgccgccatcggcgcgcccttcggcttcgac 480

Qy     634 aacagcgtgaccgccggcatcgtgtccgccaaaggcagaagcctgcccaacgaaagctac 693
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     481 aacagcgtgaccgccggcntcgtgtccgccaaaggcagaagcctgcccaacgaaagctac 540

Qy     694 acacccttcatccaaaccgacgttgccatcaatccgggcaattccggcggcccgctgttc 753
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     541 acacccttcatccaaaccgacgttgccatcaatccgggcaactccggcggcccgctgttc 600

Qy     754 aacttaaaaggacaggtcgtcggcatcaattcgcaaatatacagccgcagcggcggattc 813
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     601 aacttaaaaggacaggtcgtcggcatcaactcgcaaatatacagccgcagcggcggattc 660

Qy     814 atgggcatctcctttgccatcccgattgacgttgccatgaatgtcgccgaacagctgaaa 873
           |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     661 atgggcatttccttcgccatcccgattgacgttgccatgaatgtcgccgaacagctgaaa 720

Qy     874 aacaccggcaaagtccaacgcggacaactgggcgtgattattcaggaagtatcctacggt 933
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     721 aacaccggcaaagtccaacgcggacaactgggcgtgattattcaagaagtatcctacggt 780

Qy     934 ttggcacagtcgttcggtctggataaagccagcggcgcattgattgccaaaatccttccc 993
           |||||||| |||||||| |||| |||||| |||||||| |||||||||||||||| |||
Db     781 ttggcacaatcgttcggtttggacaaagccggcggcgcactgattgccaaaatcctgccc 840

Qy     994 ggcagccccgcagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgacggc 1053
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     841 ggcagccccgcagaacgtgccggcctgcgggcgggcgacatcgtcctcagcctcgacggc 900

Qy     1054 ggagaaatacgttcttccggcgaccttcccgtcatggtcggcgccattacgccgggaaaa 1113
            |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     901 ggagaaatacgttcttccggcgaccttcccgttatggtcggcgccattacgccgggaaaa 960

Qy     1114 gaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagccaagctgggc 1173
            |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     961 gaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagtcaagctgggc 1020

Qy     1174 aacgccgccgagcataccggcgcatcatccaaaacagatgaagccccctacaccgaacag 1233
            |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     1021 aacgccgccgagcatatcggcgcatcatccaaaacagatgaagccccctacaccgaacag 1080

Qy     1234 caatccggtacgttctcggtcgaatccgcaggcattacccttcagacacataccgacagc 1293
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 caatccggtacgttctcggtcgaatccgcaggcattacccttcagacacataccgacagc 1140

Qy     1294 agcggcaaacacctcgtcgtcgtacgggtttccgacgcggcagaacgcgcaggcttaagg 1353
            |||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||
Db     1141 agcggcggacacctcgtcgtcgtacgggtttccgacgcggcagaacgcgcaggcttgagg 1200

Qy     1354 cacggcgacgaaatcctagccgtcagggcaagtccccgtcaa 1395
            | |||||||||||| || |||||| ||||||||||||||||
Db     1201 cgcggcgacgaaattcttgccgtc-gggcaagtccccgtcaa 1241
```

```
RESULT  12
Z33308
ID   Z33308 standard; DNA; 1110 BP.
XX
AC   Z33308;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Neisseria meningitidis BASB013-C nucleotide sequence.
XX
KW   Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW   antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW   invasive bacterial disease; antibacterial; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9955872-A1.
XX
PD   04-NOV-1999.
XX
PF   20-APR-1999;   99WO-EP02765.
XX
PR   23-APR-1998;   98GB-0008734.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-052809/04.
DR   P-PSDB; Y52996.
XX
PT   Novel polynucleotides and polypeptides from Neisseria meningitis used
PT   to prepare vaccines against bacterial infections  -
XX
PS   Example 2; Page 81; 94pp; English.
XX
CC   The present sequence encodes a conserved BASB013-C polypeptide isolated
CC   from Neisseria meningitidis. BASB013 polynucleotides and polypeptides
CC   may be employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence 1110 BP; 280 A; 350 C; 279 G; 201 T; 0 other;

  Query Match            76.1%;  Score 1061;  DB 21;  Length 1110;
  Best Local Similarity  97.3%;  Pred. No. 6.3e-278;
  Matches 1079;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy       2 tgttcaaaaaataccaatacttcgctttggcggcactgtgtgccgccttgctggcaggct 61
           |||||||||||||||||||| |||||||||| |||||||||||||||| ||||||||||
Db       2 tgttcaaaaaataccaatacctcgctttggcagcactgtgtgccgcctcgctggcaggct 61

Qy      62 gcgaaaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgca 121
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 gcgacaaggcaggcagctttttcggtgcggacaaaaaagaagcatccttcgtagaacgca 121

Qy     122 tcgaacacaccaaagacgacggcagtgtcagtatgctgctgcccgactttgcccaact
           ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     122 tcgaacacaccaaagacgacggcagcgtcagtatgctgctgcccgactttgcccaactgg aaa

Qy     182 ttcaaagcgaaggcccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaaa 241
           ||||||| ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     182 ttcaaagtgaaggtccggcagtcgtcaatattcaggcagcccccgccccgcgcacccaaa 241

Qy     242 acggcagcggcaatgccgaaaccgattccgacccgcttgccgacagcgacccgttctacg 301
           |||||||||||||||||||| |||||||||||| ||||||||| ||||||||||||||
Db     242 acggcagcggcaatgccgaaaacgattccgacccgattgccgacaacgacccgttctacg 301

Qy     302 aatttttcaaacgcctcgtcccgaacatgcccgaaatcccccaagaagaagcagatgacg 361
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     302 aatttttcaaacgcctcgtcccgaatatgcccgaaatcccccaagaagaagcagatgacg 361

Qy     362 gcggattgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaata 421
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     362 gcggattgaacttcggttcgggcttcatcatcagcaaagacggctacatcctgaccaata 421

Qy     422 cccacgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatata 481
           |||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db     422 cccacgtcgttaccggcatgggcagtatcaaagtcctgctcaacgacaagcgcgaatata 481

Qy     482 ccgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaa 541
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 ccgccaaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaa 541

Qy     542 cggaagagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatggg 601
           |||||||||| |||||||||||||||||||||||||| |||||||||||||||||||||
Db     542 cggaagagctgcccgtcgtcaaaatcggcaatcccaaagatttgaaaccgggcgaatggg 601

Qy     602 tcgctgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccg 661
           |||| ||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     602 tcgccgccatcggcgcgcccttcggcttcgacaacagcgtgaccgccggcatcgtgtccg 661

Qy     662 ccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcca 721
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 ccaaaggcagaagcctgcccaacgaaagctacacacccttcatccaaaccgacgttgcca 721

Qy     722 tcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatca 781
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     722 tcaatccgggcaactccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatca 781

Qy     782 attcgcaaatatacagccgcagcggcggattcatgggcatctccttgccatcccgattg 841
           | |||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db     782 actcgcaaatatacagccgcagcggcggattcatgggcatttccttcgccatcccgattg 841

Qy     842 acgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaac 901
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 acgttgccatgaatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaac 901

Qy     902 tgggcgtgattattcaggaagtatcctacggtttggcacagtcgttcggtctggataaag 961
           |||||||||||||||| |||||||||||||||||||||| ||||||| |||| ||||
Db     902 tgggcgtgattattcaagaagtatcctacggtttggcacaatcgttcggtttggacaaag 961

Qy     962 ccagcggcgcattgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgc 1021
           || |||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db     962 ccggcggcgcactgattgccaaaatcctgcccggcagccccgcagaacgtgccggcctgc 1021

Qy    1022 aggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttc 1081
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1022 aggcgggcgacatcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttc 1081

Qy    1082 ccgtcatggtcggcgccattacgccggga 1110
           |||| |||||||||||||||||||||||
Db    1082 ccgttatggtcggcgccattacgccggga 1110


RESULT  13
X98241
ID   X98241 standard; DNA; 1436 BP.
```

```
XX
AC   X98241;
XX
DT   25-OCT-1999  (first entry)
XX
DE   Nucleotide sequence of the PA14 degP gene.
XX
KW   Human pathogen; virulence polypeptide; virulence factor;
KW   pathogenic infection; Pseudomonas aeruginosa infection; ss.
XX
OS   Pseudomonas aeruginosa.
XX
PN   WO9927129-A1.
XX
PD   03-JUN-1999.
XX
PF   25-NOV-1998;   98WO-US25247.
XX
PR   25-NOV-1997;   97US-0066517.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Ausubel F,  Cao H,  Drenkard E,  Goodman HM,  Mahajan-Miklos S;
PI   Rahme LG,  Tan M,  Tsongalis J;
XX
DR   WPI; 1999-357851/30.
XX
PT   Virulence factors useful in developing disease treatments
XX
PS   Disclosure; Fig 27; 228pp; English.
XX
CC   The present sequence represents a Pseudomonas aeruginosa nucleic aicd
CC   sequence. P. aeruginosa is an opportunistic human pathogen present in
CC   soil water and plants. The specification describes virulence polypeptides
CC   and nucleic acid sequence encoding such polypeptides. These sequences
CC   can be used to identify a compound which is capable of decreasing the
CC   expression of a pathogenic virulence factor. Compounds that inhibit
CC   the expression or activity of virulence factor polypeptides can be
CC   used to treat pathogenic infections, especially where the infection
CC   is a P. aeruginosa infection.
CC   note: the sequences given in the specification were poorly legible, and
CC   in some instances assumptions were made as to the identity of the base;
CC   it is therefore possible that the sequence given below is not entirely
CC   correct.
XX
SQ   Sequence 1436 BP; 267 A; 461 C; 471 G; 237 T; 0 other;

  Query Match             26.2%;  Score 364.8;  DB 20;  Length 1436;
  Best Local Similarity   66.0%;  Pred. No. 2.1e-89;
  Matches  544;  Conservative    0;  Mismatches  277;  Indels    3;  Gaps    1;

Qy    374 tcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttg 433
          | || || ||||||||||||  |||  ||||||||||||||| ||||| |  |||||||| |
Db    346 tgggctccggcttcatcatctccaacgacggctacatcctcaccaacaatcacgtcgtgg 405

Qy    434 ccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactca 493
          |||  |   | |   |||   || |  ||   ||||    ||||  | |  ||||| ||
Db    406 ccgatgccgacgagatcctggtgcgcctgtccgaccgtagcgagcacaaggccaagctgg 465

Qy    494 tcggttcggatgtccaatccgatgtcgccccttctgaaaatcgacgcaacggaagagctac 553
          ||||  ||||    |    ||| || ||  | ||||| ||||| |   || |  | || |
Db    466 tcggcgcggacccgcgcagcgacgtggcggtgctgaagatcgagg---cgaagaacctgc 522

Qy    554 ccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcg 613
          |   | | ||| | ||| || | || ||  |||||  ||||||||||||   |||||||
Db    523 cgaccctgaaactgggcgattcgaacaagctgaaagtgggcgaatgggtcctggccatcg 582

Qy    614 gcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcagaa 673
          |  |||| |||||||| ||  ||   || |||||||| |||||    ||||| ||  | |
Db    583 gttcgccgttcggcttcgatcactcggtcaccgccggtatcgtcagtgccaaggggcgta 642

Qy    674 gcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggca 733
          ||||||| |||||  ||||||  |||||||||||| |||||||| || |||||  ||||||| 
Db    643 gcctgccgaacgagagctacgtacccttcatccagaccgacgtggcgatcaacccgggca 702

Qy    734 attccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatat 793
          | |||||||| ||||||  | |||  |   |  ||   | ||  ||||||||||| ||||  || |
Db    703 actccggcggtccgctgctgaacctggagggcgaagtggtcggcatcaactcgcagatct 762

Qy    794 acagccgcagcggcggattcatgggcatctccttttgccatcccgattgacgttgccatga 853
            || |||   ||||||  |||||||||  | |||||  ||||||||||| || || ||  |||
Db    763 tcacccgttccggcggcttcatgggcctgtccttcgccatcccgatcgatgtcgcgctga 822

Qy    854 atgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgatta 913
          | |||||||| ||| |||| ||  |||||||  |||   |||||    ||||| ||| | |
Db    823 acgtcgccgaccagttgaagaaagccggcaaggtcagccgcggctggctgggtgtggtga 882

Qy    914 ttcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcat 973
          | ||||||||   | | | | | ||  |||| ||||| || || ||  |   ||||||
Db    883 tccaggaagtgaacaaggatctcgccgagtccttcggcctcgacaagccgtccggcgcgc 942

Qy    974 tgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgaca 1033
          || | ||  |  |  |    | | | || || |     | ||||||||||| ||||||
Db    943 tggtggcgcagctggtggaagacggtccggcggccaagggcggcctgcaggtgggcgatg 1002

Qy    1034 tcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcg 1093
           | ||||||||||  |||||      ||       |||| ||||||  ||   | |||| |
Db    1003 tgatcctcagcctgaacggccagtcgatcaacgagtccgccgacctgccgcacctggtgg 1062

Qy    1094 gcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatca 1153
           ||  ||| | ||||||  |  |  ||| ||| | |||    |||||  ||| |    |
Db    1063 gcaacatgaagccgggcgacaagatcaacctggacgtgattcgcaacggccagcgcaagt 1122

Qy    1154 caatcaaagccaagctgggcaacgccgccgagcataccggcgca 1197
           |  | |       | | ||||||    | ||  |   ||  | |
Db    1123 ccttgagcatggcggtaggcaaccttccggacgacgacgaggaa 1166


RESULT  14
Z33309
ID   Z33309 standard; DNA; 390 BP.
XX
AC   Z33309;
XX
DT   21-FEB-2000  (first entry)
XX
DE   Neisseria meningitidis variable BASB013-V nucleotide sequence.
XX
KW   Neisseria meningitidis; BASB013; diagnosis; infection; vaccine;
KW   antibiotic; upper respiratory tract infection; bacteraemia; meningitis;
KW   invasive bacterial disease; antibacterial; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9955872-A1.
XX
PD   04-NOV-1999.
XX
PF   20-APR-1999;   99WO-EP02765.
XX
PR   23-APR-1998;   98GB-0008734.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-052809/04.
DR   P-PSDB; Y52997.
XX
PT   Novel polynucleotides and polypeptides from Neisseria meningitis used
PT   to prepare vaccines against bacterial infections
XX
PS   Example 2; Page 83; 94pp; English.
XX
CC   The present sequence encodes a variable BASB013-V polypeptide isolated
```

```
CC   from Neisseria meningitidis. BASB013 polynucleotides and polypeptides
CC   may be employed as research reagents and material for the discovery of
CC   treatments and diagnostics for diseases, particularly human diseases.
CC   They can be used for diagnosis of disease, staging of disease, or
CC   determining response of an infectious organism to drugs. The
CC   polynucleotides may be used as a source for hybridisation probes, and
CC   for screening of genetic mutations, serotype, organism or strain
CC   identification, identification of mutation in BASB013 sequences, and as
CC   components of arrays which are useful for diagnostic and prognostic
CC   purposes. The polypeptides can be used to produce antibodies. The
CC   polypeptides can also be used in vaccine formulations, and to identify
CC   agonists and antagonists. The polypeptides, antibodies, agonists and
CC   antagonists (which are bacteristatic) are used for the treatment and
CC   prevention of diseases such as upper respiratory tract infection,
CC   invasive bacterial diseases such as bacteraemia and meningitis, and for
CC   the development and screening of antibacterial drugs. They are also used
CC   in the prevention of adhesion of bacteria to eukaryotic matrix proteins
CC   on in-dwelling devices, or to extracellular proteins on wounds, and to
CC   thus prevent tissue damage and/or block the normal progression of
CC   pathogenesis in infections initiated other than by the implantation of
CC   in-dwelling devices or by other surgical techniques.
XX
SQ   Sequence 390 BP; 106 A; 120 C; 102 G; 62 T; 0 other;


  Query Match             18.7%;  Score 260.2;  DB 21;  Length 390;
  Best Local Similarity   96.8%;  Pred. No. 2.8e-61;
  Matches  276;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy   1111 aaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagccaagctg 1170
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db      1 aaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagtcaagctg 60

Qy   1171 ggcaacgccgccgagcataccggcgcatcatccaaaacagatgaagccccctacaccgaa 1230
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     61 ggcaacgccgccgagcatatcggcgcatcatccaaaacagatgaagccccctacaccgaa 120

Qy   1231 cagcaatccggtacgttctcggtcgaatccgcaggcattacccttcagacacataccgac 1290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 cagcaatccggtacgttctcggtcgaatccgcaggcattacccttcagacacataccgac 180

Qy   1291 agcagcggcaaacacctcgtcgtcgtacgggtttccgacgcggcagaacgcgcaggctta 1350
           |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 agcagcggcggacacctcgtcgtcgtacgggtttccgacgcggcagaacgcgcaggcttg 240

Qy   1351 aggcacggcgacgaaatcctagccgtcagggcaagtccccgtcaa 1395
           |||| |||||||||||| || |||||| ||||||||||||||||
Db    241 aggcgcggcgacgaaattcttgccgtc-gggcaagtccccgtcaa 284


RESULT  15
V29528
ID   V29528 standard; cDNA; 1854 BP.
XX
AC   V29528;
XX
DT   13-OCT-1998  (first entry)
XX
DE   Escherichia coli serine protease htrA.
XX
KW   PS-1; presenilin; presenilin-1; PSP-1; Alzheimer's disease; htrA;
KW   serine protease; neurodegeneration; predisposition; diagnosis; ss.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   CDS             199..1674
FT                   /*tag= a
FT                   /product= htrA serine protease
XX
PN   EP828003-A2.
XX
PD   11-MAR-1998.
XX
PF   26-AUG-1997;   97EP-0306501.
XX
PR   13-DEC-1996;   96US-0032875.
PR   06-SEP-1996;   96US-0025436.
PR   25-OCT-1996;   96US-0027873.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Browne MJ,  Clinkenbeard HE,  Creasy CL,  Karran EH;
PI   Livi GP,  Southan CD;
XX
DR   WPI; 1998-161101/15.
DR   P-PSDB; W56771.
XX
PT   Nucleic acids encoding human serum protease protein(s) - used for
PT   diagnosing pre-disposition to Alzheimer's disease, etc.
XX
PS   Example 2; Page 27-28; 65pp; English.
XX
CC   The sequence is that encoding the serine protease htrA which was used
CC   in the isolation of PSP1.
XX
SQ   Sequence 1854 BP; 478 A; 470 C; 489 G; 417 T; 0 other;


  Query Match             12.6%;  Score 176.4;  DB 19;  Length 1854;
  Best Local Similarity   52.0%;  Pred. No. 2.4e-38;
  Matches  421;  Conservative    0;  Mismatches  386;  Indels    3;  Gaps    1;

Qy    382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
          | | ||||      | | || |||||  || | ||||| | |||||| ||||       |
Db    547 gtcatcattgatgccgataaaggctatgtcgtcaccaacaaccacgttgttgataacgcg 606

Qy    442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
              || ||||| |  || | |||    ||  | |    ||| ||  |  | ||
Db    607 acggtcattaaagttcaactgagcgatggccgtaagttcgacgcgaagatggttggcaaa 666

Qy    502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
          |||   |  || ||| |||| ||  |  ||||| |     || || | ||  |||   |
Db    667 gatccgcgctctgatatcgcgctgatccaaatccagaacccgaaaaacctgaccgcaatt 726

Qy    562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
          || || |   || |  |     ||     ||| || |    ||  |  || ||    ||
Db    727 aagatggcggattctgatgcactgcgcgtgggtgattacaccgtagggattggtaacccg 786

Qy    622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaag---gcagaagcctg 678
          || ||  | | | | |  || ||  |||| || || || ||    |   | ||  |||||
Db    787 tttggtctgggcgagacggtaacttccgggattgtctctgcgctggggcgtagcggcctg 846

Qy    679 cccaacgaaagctacacaccttcatccaaaccgacgttgccatcaatccgggcaattcc 738
              ||||| ||||  |  |||||||| ||||| |  || ||||| |  || || |||
Db    847 aatgccgaaaactacgaaaacttcatccagaccgatgcagcgatcaaccgtggtaactcc 906

Qy    739 ggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatacagc 798
          || ||  ||||| | ||| | || ||  |  |  |||| |||||  |    ||  |
Db    907 ggtggtgcgctggttaacctgaacggcgaactgatcggtatcaacaccgcgatcctcgca 966

Qy    799 cgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaatgtc 858
          |   ||||||   ||| || |||   ||||| ||||||| | || | |  |  ||  |
Db    967 ccggacggcggcaacatcggtatcggttttgctatcccgagtaacatggtgaaaaacctg 1026

Qy    859 gccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattattcag 918
           ||   ||| ||    |   |||| | ||  |||||||  | |||||  | ||
Db   1027 acctcgcagatggtggaatacggccaggtgaaacgcggtgagctgggtattatggggact 1086

Qy    919 gaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattgatt 978
          ||  |   || ||   ||||  |  || |     | ||        |||| || ||  |
Db   1087 gagctgaactccgaactggcgaaagcgatgaaagttgacgcccagcgcggtgctttcgta 1146

Qy    979 gccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatcgtc 1038
            | |  | || ||      | |||| | |   || ||| |  | |||||  ||  |  ||
```

```
Db   1147 agccaggttctgcctaattcctccgctgcaaaagcgggcattaaagcgggtgatgtgatc 1206

Qy   1039 ctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggcgcc 1098
            |   ||  ||||       ||  |    |  | ||  || |  |    ||| ||  |
Db   1207 acctcactgaacggtaagccgatcagcagctttgccgcactgcgtgctcaggtgggtact 1266

Qy   1099 attacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatc 1158
          ||  ||   || |   || | | ||| ||| ||  |||| | ||  |  |  | |   |
Db   1267 atgccggtaggcagcaaactgaccctgggcttactgcgcgacggtaagcaggttaacgtg 1326

Qy   1159 aaagccaagctgggcaacgccgccgagcat 1188
          ||     | |||    |   |  | || ||
Db   1327 aacctggaactgccgcagagcagccagaat 1356


Search completed: March 22, 2001, 10:17:11
Job time: 25016 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 02:51:15 ; Search time 58.35 Seconds
(without alignments)
3852.933 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaaataccaata.....tcagggcaagtccccgtcaa 1395

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	176.4	12.6	1854	3	US-08-923-454A-13	Sequence 13, Appl
2	175.6	12.6	1980	1	US-08-350-741-1	Sequence 1, Appli
3	175.6	12.6	1980	2	US-08-463-875A-1	Sequence 1, Appli
4	154.2	11.1	1791	1	US-08-245-294-7	Sequence 7, Appli
5	154.2	11.1	1791	1	US-08-474-499-7	Sequence 7, Appli
6	154.2	11.1	1791	1	US-08-307-279A-7	Sequence 7, Appli
7	154.2	11.1	1791	4	PCT-US95-06211-7	Sequence 7, Appli
8	139.4	10.0	2894	1	US-08-278-091-1	Sequence 1, Appli
9	139.4	10.0	2894	1	US-08-483-859-1	Sequence 1, Appli
10	139.4	10.0	2894	1	US-08-472-173-1	Sequence 1, Appli
11	139.4	10.0	2894	2	US-08-487-167-1	Sequence 1, Appli
12	139.4	10.0	2894	2	US-08-482-816-1	Sequence 1, Appli
13	139.4	10.0	2894	2	US-08-296-149-1	Sequence 1, Appli
14	139.4	10.0	2894	2	US-08-801-499-1	Sequence 1, Appli
15	139.4	10.0	2894	2	US-08-615-271-1	Sequence 1, Appli
16	139.4	10.0	2894	3	US-09-074-660-1	Sequence 1, Appli
17	139.4	10.0	2894	3	US-09-074-659-1	Sequence 1, Appli
18	139.4	10.0	2894	3	US-09-106-468-1	Sequence 1, Appli
19	139.4	10.0	2894	3	US-09-106-466A-1	Sequence 1, Appli
20	139.4	10.0	2894	3	US-09-106-467-1	Sequence 1, Appli
21	136.2	9.8	1611	1	US-08-485-569-1	Sequence 1, Appli
22	136.2	9.8	1611	1	US-08-480-993-1	Sequence 1, Appli
23	136.2	9.8	1611	2	US-07-903-079B-1	Sequence 1, Appli
24	105	7.5	2036	3	US-08-923-454A-17	Sequence 17, Appl
25	105	7.5	2205	3	US-08-888-077A-41	Sequence 41, Appl
26	85.8	6.2	539	1	US-08-322-742-18	Sequence 18, Appl
27	57.6	4.1	1835	3	US-08-923-454A-7	Sequence 7, Appli
28	57.6	4.1	2187	3	US-08-923-454A-23	Sequence 23, Appl
29	57.6	4.1	2187	3	US-08-923-454A-24	Sequence 24, Appl
30	57.6	4.1	2187	3	US-08-923-454A-30	Sequence 30, Appl
31	51.6	3.7	1503	3	US-08-923-454A-5	Sequence 5, Appli
32	51.6	3.7	1787	3	US-08-923-454A-3	Sequence 3, Appli
33	51.6	3.7	2144	3	US-08-923-454A-28	Sequence 28, Appl
34	51.6	3.7	2551	3	US-08-923-454A-26	Sequence 26, Appl
35	41	2.9	1187	1	US-08-440-856A-2	Sequence 2, Appli
36	39.2	2.8	2414	5	5248599-1	Patent No. 5248599
c 37	37.2	2.7	2793	1	US-08-209-747-1	Sequence 1, Appli
c 38	37.2	2.7	2793	1	US-08-458-298-1	Sequence 1, Appli
c 39	36.8	2.6	716	2	US-08-211-718-1	Sequence 1, Appli
c 40	36.8	2.6	1608	2	US-08-211-718-8	Sequence 8, Appli
c 41	36.2	2.6	411	3	US-08-963-168C-1	Sequence 1, Appli
c 42	36.2	2.6	411	3	US-08-963-168C-2	Sequence 2, Appli
c 43	35.8	2.6	399	3	US-08-963-168C-5	Sequence 5, Appli
44	35.8	2.6	2167	2	US-08-461-775-9	Sequence 9, Appli
45	35.8	2.6	2167	3	US-09-031-606-9	Sequence 9, Appli

ALIGNMENTS

```
RESULT 1
US-08-923-454A-13
; Sequence 13, Application US/08923454A
; Patent No. 6004794
;  GENERAL INFORMATION:
;    APPLICANT:  Creasy, Caretha
;    APPLICANT:  Livi, George
;    APPLICANT:  Karran, Eric
;    APPLICANT:  Clinkenbeard, Helen
;    APPLICANT:  Browne, Michael
;    APPLICANT:  Southan, Christopher
;    TITLE OF INVENTION:  HUMAN SERINE PROTEASE
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/923,454A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/025436
;      FILING DATE:  06-SEPT-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Baumeister, Kirk
;      REGISTRATION NUMBER:  33,833
;      REFERENCE/DOCKET NUMBER:  P50547
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5096
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1854 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
```

US-08-923-454A-13

Query Match 12.6%; Score 176.4; DB 3; Length 1854;
Best Local Similarity 52.0%; Pred. No. 1.5e-39;
Matches 421; Conservative 0; Mismatches 386; Indels 3; Gaps 1;

```
Qy    382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
          | | ||||       | | || |||||  || | ||||| | |||||| ||||       |
Db    547 GTCATCATTGATGCCGATAAAGGCTATGTCGTCACCAACAACCACGTTGTTGATAACGCG 606

Qy    442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
              || ||||| |  || | |||    ||  | |    ||| ||  |  | ||
Db    607 ACGGTCATTAAAGTTCAACTGAGCGATGGCCGTAAGTTCGACGCGAAGATGGTTGGCAAA 666

Qy    502 gatgtccaatccgatgtcgccccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
          |||   |  || ||| |||| ||  |  ||||| |     || || | ||  |||   |
Db    667 GATCCGCGCTCTGATATCGCGCTGATCCAAATCCAGAACCCGAAAAACCTGACCGCAATT 726

Qy    562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
          || || |   || |  |     ||     ||| || |    ||  |  || ||    ||
Db    727 AAGATGGCGGATTCTGATGCACTGCGCGTGGGTGATTACACCGTAGGGATTGGTAACCCG 786

Qy    622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaag---gcagaagcctg 678
          || ||  | | | | |  || ||  |||| || || || ||    |   | ||  |||||
Db    787 TTTGGTCTGGGCGAGACGGTAACTTCCGGGATTGTCTCTGCGCTGGGGCGTAGCGGCCTG 846

Qy    679 cccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaattcc 738
              ||||| ||||  |  |||||||||| ||||| |  || ||||| |  || || |||
Db    847 AATGCCGAAAACTACGAAAACTTCATCCAGACCGATGCAGCGATCAACCGTGGTAACTCC 906

Qy    739 ggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatacagc 798
          || ||  ||||| | ||| | || ||  |  |  |||| |||||  |    ||   |
Db    907 GGTGGTGCGCTGGTTAACCTGAACGGCGAACTGATCGGTATCAACACCGCGATCCTCGCA 966

Qy    799 cgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaatgtc 858
          |    ||||||   ||| || |||   ||||| ||||||| | || | |  | ||  |
Db    967 CCGGACGGCGGCAACATCGGTATCGGTTTTGCTATCCCGAGTAACATGGTGAAAAACCTG 1026

Qy    859 gccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattattcag 918
            ||   ||| ||    |   |||| | ||  |||||||  | |||||  | ||
Db   1027 ACCTCGCAGATGGTGGAATACGGCCAGGTGAAACGCGGTGAGCTGGGTATTATGGGGACT 1086

Qy    919 gaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattgatt 978
          ||  |   || ||   ||||  |  || |     | ||        |||| || ||  |
Db   1087 GAGCTGAACTCCGAACTGGCGAAAGCGATGAAAGTTGACGCCCAGCGCGGTGCTTTCGTA 1146

Qy    979 gccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatcgtc 1038
            | |  | || ||      | |||| | |   || ||| |  | ||||| ||  |  ||
Db   1147 AGCCAGGTTCTGCCTAATTCCTCCGCTGCAAAAGCGGGCATTAAAGCGGGTGATGTGATC 1206

Qy    1039 ctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggcgcc 1098
            |  ||  ||||      ||  |    |  | ||  || |  |    ||| ||  |
Db   1207 ACCTCACTGAACGGTAAGCCGATCAGCAGCTTTGCCGCACTGCGTGCTCAGGTGGGTACT 1266

Qy    1099 attacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatc 1158
          ||  ||   || |   || | | ||| ||| ||  |||| | ||  |  |  | |   |
Db   1267 ATGCCGGTAGGCAGCAAACTGACCCTGGGCTTACTGCGCGACGGTAAGCAGGTTAACGTG 1326

Qy    1159 aaagccaagctgggcaacgccgccgagcat 1188
          ||     | |||    |   |  | || ||
Db   1327 AACCTGGAACTGCAGCAGAGCAGCCAGAAT 1356
```

RESULT 2
US-08-350-741-1
; Sequence 1, Application US/08350741
; Patent No. 5804194
; GENERAL INFORMATION:
; APPLICANT: DOUGAN G.,
; APPLICANT: CHARLES I.G.,
; APPLICANT: HORMAECHE C.E.,
; APPLICANT: JOHNSON K.S.,
; APPLICANT: CHATFIELD S.N.
; TITLE OF INVENTION: LIVE VACCINES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON and VANDERHYE PC
; STREET: 8th FLOOR, 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/350,741
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952,737
; FILING DATE: 09-MAY-1994
; APPLICATION NUMBER: US 07/952,737
; FILING DATE: 30-NOV-1992
; APPLICATION NUMBER: GB 9007194.5
; FILING DATE: 30-MAR-1990
; APPLICATION NUMBER: PCT/GB91/00484
; FILING DATE: 28-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-158
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 395..1822
US-08-350-741-1

Query Match 12.6%; Score 175.6; DB 1; Length 1980;
Best Local Similarity 52.6%; Pred. No. 2.6e-39;
Matches 407; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

```
Qy    400 aacggctacatcctgaccaatacccacgtcgttgccggtatgggcagtatcaaagtcctg 459
          || |||||| || | ||||| | |||||| ||||          ||   || ||||| | |
Db    764 AAGGGCTACGTCGTCACCAACAACCACGTGGTTGATAACGCCAGCGTGATTAAAGTACAG 823

Qy    460 ctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatccgatgtc 519
          || | |||   |||  |||      || ||| |  | ||    |||   |  || ||| ||
Db    824 CTTAGCGATGGGCGTAAATTCGATGCTAAAGTGGTGGGCAAAGATCCGCGTTCTGATATC 883

Qy    520 gcccttctgaaaatcgacgcaacggaagagctacccgtcgtcaaaatcggcaatcccaaa 579
          || ||  |  ||||  |     || |  | ||  | |   | ||  | |   |  || |
Db    884 GCGCTGATTCAAATTCAGAATCCGAAGAACCTGACGGCGATTAAGCTGGCGGACTCCGAC 943

Qy    580 aatttgaaaccgggcgaatgggtcgctgccatcggcgcgcccttcggctttgacaacagc 639
              ||     ||| || |   ||  || || ||    || || ||  | | | | |
Db    944 GCGCTGCGCGTGGGGGATTATACCGTCGCTATTGGTAACCCGTTTGGTCTGGGCGAAACG 1003

Qy    640 gtgaccgccggcatcgtgtccgccaaag---gcagaagcctgcccaacgaaagctacaca 696
          |||||  | || ||||| || ||    |   | ||  |||||  |   |||| |||
Db   1004 GTGACGTCAGGTATCGTTTCGGCGCTGGGGCGTAGCGGCCTGAACGTAGAAAATTACGAG 1063
```

```
Qy     697 cccttcatccaaaccgacgttgccatcaatccgggcaattccggcggcccgctgttcaac 756
           ||| || || |||||||  || || || |  || || |||||||||| ||||| | |||
Db    1064 AACTTTATTCAGACCGACGCCGCGATTAACCGTGGTAACTCCGGCGGCGCGCTGGTGAAC 1123

Qy     757 ttaaaaggacaggtcgtcggcatcaattcgcaaatatacagccgcagcggcggattcatg 816
           | || ||  || |  |||| || ||  |    ||       |    ||||||   |||
Db    1124 CTGAACGGTGAGCTGATCGGTATTAACACCGCGATTCTGGCGCCGGACGGCGGCAACATC 1183

Qy     817 ggcatctcctttgccatcccgattgacgttgccatgaatgtcgccgaacagctgaaaaac 876
           || |||  ||| || ||||| | | || | |  |  ||  |  |    ||| ||    |
Db    1184 GGTATCGGCTTCGCTATCCCCAGTAACATGGTGAAAAACCTGACGTCGCAGATGGTGGAA 1243

Qy     877 accggcaaagtccaacgcggacaactgggcgtgattattcaggaagtatcctacggtttg 936
             |||| | ||  |||||||  |||||||  | ||       ||  |    | ||  |||
Db    1244 TACGGCCAGGTGAAACGCGGCGAACTGGGGATCATGGGGACTGAGCTGAATTCCGAATTG 1303

Qy     937 gcacagtcgttcggtctggataaagccagcggcgcattgattgccaaaatccttcccggc 996
           ||  |  || |     | ||         | ||||| ||  |   | |  |  | ||
Db    1304 GCGAAAGCGATGAAAGTCGACGCCCAGCGAGGCGCGTTCGTCAGCCAGGTGATGCCGAAT 1363

Qy     997 agccccgcagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgacggcgga 1056
           ||||  |     || ||  |  | || || ||  || |   |   ||  |||| |
Db    1364 TCGTCCGCGGCGAAAGCGGGTATCAAAGCCGGGGATGTCATTACCTCGCTGAACGGTAAA 1423

Qy    1057 gaaatacgttcttccggcgaccttcccgtcatggtcggcgccattacgccgggaaaagaa 1116
             ||  |    |  |  || | ||    |||||||  | ||  ||   || |   ||
Db    1424 CCGATCAGCAGCTTTGCGGCGCTGCGCGCTCAGGTCGGCACTATGCCGGTCGGCAGCAAA 1483

Qy    1117 gtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagccaagctg 1170
           ||||||||||  |   |||  ||||  |||  || ||  | ||     | |||
Db    1484 ATCAGCCTCGGTCTGCTGCGTGAAGGTAAAGCGATTACGGTGAATCTGGAACTG 1537
```

RESULT 3
US-08-463-875A-1
; Sequence 1, Application US/08463875A
; Patent No. 5980907
; GENERAL INFORMATION:
; APPLICANT: DOUGAN, Gordon
; APPLICANT: CHARLES, Ian G.
; APPLICANT: HORMAECHE, Carlos E.
; APPLICANT: JOHNSON, Kevin S.
; APPLICANT: CHATFIELD, Steven N.
; TITLE OF INVENTION: LIVE VACCINES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON and VANDERHYE PC
; STREET: 8th FLOOR, 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,875A
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/340,741
; FILING DATE: 07-DEC-1994
; APPLICATION NUMBER: US 07/952,737
; FILING DATE: 30-NOV-1992
; APPLICATION NUMBER: GB 9007194.5
; FILING DATE: 30-MAR-1990
; APPLICATION NUMBER: PCT/GB91/00484
; FILING DATE: 28-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 117-158
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 395..1822
US-08-463-875A-1

Query Match 12.6%; Score 175.6; DB 2; Length 1980;
Best Local Similarity 52.6%; Pred. No. 2.6e-39;
Matches 407; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

```
Qy     400 aacggctacatcctgaccaatacccacgtcgttgccggtatgggcagtatcaaagtcctg 459
           || |||||| || | ||||| | |||||| |||| ||  || ||||| | |
Db     764 AAGGGCTACGTCGTCACCAACAACCACGTGGTTGATAACGCCAGCGTGATTAAAGTACAG 823

Qy     460 ctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatccgatgtc 519
           || | |||   |||  |||     || ||| |  | ||   |||   |  || ||| ||
Db     824 CTTAGCGATGGGCGTAAATTCGATGCTAAAGTGGTGGGCAAAGATCCGCGTTCTGATATC 883

Qy     520 gcccttctgaaaatcgacgcaacggaagagctacccgtcgtcaaaatcggcaatcccaaa 579
           || ||  |  |||| |     || |  | ||  | |   | ||  | |   |  || |
Db     884 GCGCTGATTCAAATTCAGAATCCGAAGAACCTGACGGCGATTAAGCTGGCGGACTCCGAC 943

Qy     580 aatttgaaaccgggcgaatgggtcgctgccatcggcgcgcccttcggctttgacaacagc 639
           ||     ||| || |   ||  || || ||    || || ||  | | | | |
Db     944 GCGCTGCGCGTGGGGGATTATACCGTCGCTATTGGTAACCCGTTTGGTCTGGGCGAAACG 1003

Qy     640 gtgaccgccggcatcgtgtccgccaaag---gcagaagcctgcccaacgaaagctacaca 696
           |||||  | || ||||| || ||    |   | ||  |||||  |   ||||  |||
Db    1004 GTGACGTCAGGTATCGTTTCGGCGCTGGGGCGTAGCGGCCTGAACGTAGAAAATTACGAG 1063

Qy     697 cccttcatccaaaccgacgttgccatcaatccgggcaattccggcggcccgctgttcaac 756
           ||| || || |||||||  || || || |  || || |||||||||| ||||| | |||
Db    1064 AACTTTATTCAGACCGACGCCGCGATTAACCGTGGTAACTCCGGCGGCGCGCTGGTGAAC 1123

Qy     757 ttaaaaggacaggtcgtcggcatcaattcgcaaatatacagccgcagcggcggattcatg 816
           | || ||  || |  |||| || ||  |    ||       |    ||||||   |||
Db    1124 CTGAACGGTGAGCTGATCGGTATTAACACCGCGATTCTGGCGCCGGACGGCGGCAACATC 1183

Qy     817 ggcatctcctttgccatcccgattgacgttgccatgaatgtcgccgaacagctgaaaaac 876
           || |||  ||| || ||||| | | || | |  |  ||  |  |    ||| ||    |
Db    1184 GGTATCGGCTTCGCTATCCCCAGTAACATGGTGAAAAACCTGACGTCGCAGATGGTGGAA 1243

Qy     877 accggcaaagtccaacgcggacaactgggcgtgattattcaggaagtatcctacggtttg 936
             |||| | ||  |||||||  |||||||  | ||       ||  |    | ||  |||
Db    1244 TACGGCCAGGTGAAACGCGGCGAACTGGGGATCATGGGGACTGAGCTGAATTCCGAATTG 1303

Qy     937 gcacagtcgttcggtctggataaagccagcggcgcattgattgccaaaatccttcccggc 996
           ||  |  || |     | ||         | ||||| ||  |   | |  |  | ||
Db    1304 GCGAAAGCGATGAAAGTCGACGCCCAGCGAGGCGCGTTCGTCAGCCAGGTGATGCCGAAT 1363

Qy     997 agccccgcagaacgtgccggcctgcaggcgggcgacatcgtcctcagcctcgacggcgga 1056
           ||||  |     || ||  |  | || || ||  || |   |   ||  |||| |
Db    1364 TCGTCCGCGGCGAAAGCGGGTATCAAAGCCGGGGATGTCATTACCTCGCTGAACGGTAAA 1423

Qy    1057 gaaatacgttcttccggcgaccttcccgtcatggtcggcgccattacgccgggaaaagaa 1116
             ||  |    |  |  || | ||    |||||||  | ||  ||   || |   ||
Db    1424 CCGATCAGCAGCTTTGCGGCGCTGCGCGCTCAGGTCGGCACTATGCCGGTCGGCAGCAAA 1483

Qy    1117 gtcagcctcggcgtatggcgcaaaggcgaagaaatcacaatcaaagccaagctg 1170
           ||||||||||  |   |||  ||||  |||  || ||  | ||     | |||
Db    1484 ATCAGCCTCGGTCTGCTGCGTGAAGGTAAAGCGATTACGGTGAATCTGGAACTG 1537
```

```
RESULT   4
US-08-245-294-7
; Sequence 7, Application US/08245294
; Patent No. 5644047
;  GENERAL INFORMATION:
;    APPLICANT:  Anderson, Burt E.
;    APPLICANT:  Regnery, Russell L.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSING
;    TITLE OF INVENTION:  ROCHALIMAEA HENSELAE AND ROCHALIMAEA QUINTANA
;    TITLE OF INVENTION:  INFECTION
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NEEDLE & ROSENBERG,P.C.
;      STREET:  127 Peachtree Street, Suite 1200
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/245,294
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Spratt, Gwendolyn D.
;      REGISTRATION NUMBER:  36,016
;      REFERENCE/DOCKET NUMBER:  1414.612
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404/688-0770
;      TELEFAX:  404/688-9880
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1791 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  141..1649
US-08-245-294-7


  Query Match             11.1%;  Score 154.2;  DB 1;  Length 1791;
  Best Local Similarity   50.4%;  Pred. No. 2.1e-33;
  Matches  403;  Conservative    0;  Mismatches  393;  Indels    3;  Gaps    1;

Qy    373 ttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgtt 432
          || || ||||| ||  | |||       | || || ||  |||||||||  || ||  ||
Db    516 TTTGGATCGGGTTTTTTTATCTCGTCTGATGGTTATATTGTGACCAATAATCATGTGATT 575

Qy    433 gccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactc 492
          | | |     |||  |   ||  | ||  | |||     |||   |  || ||||||
Db    576 TCTGATGGCACAAGTTACGCTGTTGTTCTTGATGACGGTACAGAACTGAATGCAAAACTC 635

Qy    493 atcggttcggatgtccaatccgatgtcgccccttctgaaaatcgacgcaacggaagagcta 552
          || ||  ||||    | | | ||| | ||  |  | ||| || | | ||    | |  |
Db    636 ATTGGAACGGACCCACGAACTGATCTTGCAGTATTAAAAGTCAATGAAAAAAGAAAATTT 695

Qy    553 cccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatc 612
          |  |||  |  | ||  ||     |||  |       || || ||||| | ||| ||
Db    696 TCGTACGTTGATTTTGGTGATGATTCAAAACTTCGTGTTGGTGATTGGGTTGTTGCTATT 755

Qy    613 ggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcaga 672
          ||    || || ||  |||     |  ||||| || || ||||| || ||    ||  |
Db    756 GGTAATCCATTTGGTCTTGGTGGAACTGTGACAGCAGGTATCGTTTCAGCACGTGGACGT 815

Qy    673 agcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggc 732
          |    | ||     ||       || || || |  || | |||  | ||||  ||
Db    816 GATATCGGTACCGGTGTTTATGATGATTTTATTCAGATTGATGCTGCAGTTAATCGAGGA 875

Qy    733 aattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaaata 792
          ||||| || || ||    ||  |  | || ||| |||| || ||  | ||| ||  |||
Db    876 AATTCTGGAGGTCCAACTTTTGATCTTAACGGAAAGGTTGTTGGAGTGAATACGGCAATT 935

Qy    793 tacagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatg 852
          |     |    || ||   | | || ||  | || || || |||    |    || |
Db    936 TTTTCTCCTTCTGGGGGCAACGTTGGGATTGCTTTCGCTATTCCGGCAGCAACAGCGAAC 995

Qy    853 aatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgatt 912
          | || |   |||| || |   | |  ||   ||| || || ||    || || ||
Db    996 GAGGTTGTGCAACAACTTATCGAAAAAGGTTTAGTTCAGCGTGGTTGGCTTGGGGTTCAG 1055

Qy    913 attcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgca 972
          ||||||   ||| |  | |   |  |  | ||  | ||| || |  | || |  || ||
Db   1056 ATTCAGCCTGTAACAAAAGAAATTTCTGATTCAATTGGTTTGAAGGAGGCTAAAGGTGCG 1115

Qy    973 ttgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgac 1032
          || ||| || |     |    ||    ||||| ||   || ||  |  |||| || ||
Db   1116 TTAATTACCGATCCATTAAAGGGGCCAGCCGCAAAA---GCTGGTATCAAGGCAGGTGAT 1172

Qy   1033 atcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtc 1092
          |  |  |     |  | || |   | ||   |  |  |  | | || ||  |       |
Db   1173 GTTATTATTTCCGGTAAATGGTGAGAAGATTAATGATGTCCGTGATCTAGCAAAGCGTATT 1232

Qy   1093 ggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatc 1152
          |     || |  || ||| ||  ||| | | | || || |||      ||  ||||
Db   1233 GCAAATATGAGCCCAGGAGAAACAGTAACCTTAGGAGTTTGGAAATCTGGTAAAGAAGAG 1292

Qy   1153 acaatcaaagccaagctgg 1171
          |  || || |  || || |
Db   1293 AATATTAAGGTTAAACTTG 1311


RESULT   5
US-08-474-499-7
; Sequence 7, Application US/08474499
; Patent No. 5693776
;  GENERAL INFORMATION:
;    APPLICANT:  Anderson, Burt E.
;    APPLICANT:  Regnery, Russell L.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR
;    TITLE OF INVENTION:  DIAGNOSING ROCHALIMAEA HENSELAE
;    TITLE OF INVENTION:  AND ROCHALIMAEA QUINTANA INFECTION
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NEEDLE & ROSENBERG,P.C.
;      STREET:  127 Peachtree Street, Suite 1200
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,499
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/245,294
;      FILING DATE:  18-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Spratt, Gwendolyn D.
;      REGISTRATION NUMBER:  36,016
;      REFERENCE/DOCKET NUMBER:  1414.612
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404/688-0770
```

```
;      TELEFAX:  404/688-9880
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1791 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  141..1649
US-08-474-499-7


  Query Match            11.1%;  Score 154.2;  DB 1;  Length 1791;
  Best Local Similarity  50.4%;  Pred. No. 2.1e-33;
  Matches  403;  Conservative    0;  Mismatches  393;  Indels    3;  Gaps    1;

Qy    373 ttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgtt 432
          || || ||||| ||  | |||       | || || ||  |||||||||  || ||  ||
Db    516 TTTGGATCGGGTTTTTTTATCTCGTCTGATGGTTATATTGTGACCAATAATCATGTGATT 575

Qy    433 gccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactc 492
          | | |     |||  |   ||  | ||  | |||      |||   |  || ||||||
Db    576 TCTGATGGCACAAGTTACGCTGTTGTTCTTGATGACGGTACAGAACTGAATGCAAAACTC 635

Qy    493 atcggttcggatgtccaatccgatgtcgccctttctgaaaatcgacgcaacggaagagcta 552
          || ||  ||||    | | | ||| | ||  |  | ||| || | | ||   | |  |
Db    636 ATTGGAACGGACCCACGAACTGATCTTGCAGTATTAAAAGTCAATGAAAAAAGAAAATTT 695

Qy    553 cccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatc 612
          |   |||  |  | ||  ||     |||  |       || || ||||| | ||| ||
Db    696 TCGTACGTTGATTTTGGTGATGATTCAAAACTTCGTGTTGGTGATTGGGTTGTTGCTATT 755

Qy    613 ggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcaga 672
          ||    || || ||  |||     |  ||||| || || ||||| || ||    ||  |
Db    756 GGTAATCCATTTGGTCTTGGTGGAACTGTGACAGCAGGTATCGTTTCAGCACGTGGACGT 815

Qy    673 agcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggc 732
          |     | ||     ||       || || || |  || | |||  | ||||  ||
Db    816 GATATCGGTACCGGTGTTTATGATGATTTTATTCAGATTGATGCTGCAGTTAATCGAGGA 875

Qy    733 aattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaata 792
          ||||| || || ||    ||  |  | || ||| |||| || ||  | ||| ||  |||
Db    876 AATTCTGGAGGTCCAACTTTTGATCTTAACGGAAAGGTTGTTGGAGTGAATACGGCAATT 935

Qy    793 tacagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatg 852
          |     |     || ||   | | || ||  | || || || |||   |      || |
Db    936 TTTTCTCCTTCTGGGGGCAACGTTGGGATTGCTTTCGCTATTCCGGCAGCAACAGCGAAC 995

Qy    853 aatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgatt 912
          | || |   |||| || |   | |  ||   ||| || || ||    || || ||
Db    996 GAGGTTGTGCAACAACTTATCGAAAAAGGTTTAGTTCAGCGTGGTTGGCTTGGGGTTCAG 1055

Qy    913 attcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgca 972
          ||||||   ||| |  | |   |  |  | ||  | ||| || |  | || |  || ||
Db   1056 ATTCAGCCTGTAACAAAAGAAATTTCTGATTCAATTGGTTTGAAGGAGGCTAAAGGTGCG 1115

Qy    973 ttgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgac 1032
          || ||| || |     |    ||     ||||| ||   || ||  |  |||| || ||
Db   1116 TTAATTACCGATCCATTAAAGGGGCCAGCCGCAAAA---GCTGGTATCAAGGCAGGTGAT 1172

Qy   1033 atcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtc 1092
          |  |  |     |  | || |   | ||   |  |  | | || ||  |          |
Db   1173 GTTATTATTTCGGTAAATGGTGAGAAGATTAATGATGTCCGTGATCTAGCAAAGCGTATT 1232

Qy   1093 ggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatc 1152
          |     || |  || ||| ||  ||| | | | || || |||     ||  |||||
Db   1233 GCAAATATGAGCCCAGGAGAAACAGTAACCTTAGGAGTTTGGAAATCTGGTAAAGAAGAG 1292

Qy   1153 acaatcaaagccaagctgg 1171
          |  || || |  || || |
Db   1293 AATATTAAGGTTAAACTTG 1311


RESULT   6
US-08-307-279A-7
; Sequence 7, Application US/08307279A
; Patent No. 5736347
;  GENERAL INFORMATION:
;    APPLICANT:  Anderson, Burt E.
;    APPLICANT:  Regnery, Russell L
;    TITLE OF INVENTION:  Nucleic Acids of Rochalimaea Henselae
;    TITLE OF INVENTION:  and Methods and Compositions for Diagnosing Rochalimaea
;    TITLE OF INVENTION:  Henselae and Rochalimaea Quintana Infection
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NEEDLE & ROSENBERG, P.C.
;      STREET:  127 Peachtree Street, N.E., Suite 1200
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/307,279A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Spratt, Gwendolyn D.
;      REGISTRATION NUMBER:  36,016
;      REFERENCE/DOCKET NUMBER:  1414.624
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 688-0770
;      TELEFAX:  (404) 688-9880
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1791 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  141..1652
US-08-307-279A-7


  Query Match            11.1%;  Score 154.2;  DB 1;  Length 1791;
  Best Local Similarity  50.4%;  Pred. No. 2.1e-33;
  Matches  403;  Conservative    0;  Mismatches  393;  Indels    3;  Gaps    1;

Qy    373 ttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgtt 432
          || || ||||| ||  | |||       | || || ||  |||||||||  || ||  ||
Db    516 TTTGGATCGGGTTTTTTTATCTCGTCTGATGGTTATATTGTGACCAATAATCATGTGATT 575

Qy    433 gccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactc 492
          | | |     |||  |   ||  | ||  | |||      |||   |  || ||||||
Db    576 TCTGATGGCACAAGTTACGCTGTTGTTCTTGATGACGGTACAGAACTGAATGCAAAACTC 635

Qy    493 atcggttcggatgtccaatccgatgtcgccctttctgaaaatcgacgcaacggaagagcta 552
          || ||  ||||    | | | ||| | ||  |  | ||| || | | ||    | |  |
Db    636 ATTGGAACGGACCCACGAACTGATCTTGCAGTATTAAAAGTCAATGAAAAAAGAAAATTT 695

Qy    553 cccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatc 612
          |   |||  |  | ||  ||     |||  |       || || ||||| | ||| ||
Db    696 TCGTACGTTGATTTTGGTGATGATTCAAAACTTCGTGTTGGTGATTGGGTTGTTGCTATT 755

Qy    613 ggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcaga 672
          ||    || || ||  |||     |  ||||| || || ||||| || ||    ||  |
Db    756 GGTAATCCATTTGGTCTTGGTGGAACTGTGACAGCAGGTATCGTTTCAGCACGTGGACGT 815
```

```
Qy    673 agcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggc 732
Db    816 GATATCGGTACCGGTGTTTATGATGATTTTATTCAGATTGATGCTGCAGTTAATCGAGGA 875

Qy    733 aattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaata 792
Db    876 AATTCTGGAGGTCCAACTTTTGATCTTAACGGAAAGGTTGTTGGAGTGAATACGGCAATT 935

Qy    793 tacagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatg 852
Db    936 TTTTCTCCTTCTGGGGGCAACGTTGGGATTGCTTTCGCTATTCCGGCAGCAACAGCGAAC 995

Qy    853 aatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgatt 912
Db    996 GAGGTTGTGCAACAACTTATCGAAAAAGGTTTAGTTCAGCGTGGTTGGCTTGGGGTTCAG 1055

Qy    913 attcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgca 972
Db   1056 ATTCAGCCTGTAACAAAAGAAATTTCTGATTCAATTGGTTTGAAGGAGGCTAAAGGTGCG 1115

Qy    973 ttgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgac 1032
Db   1116 TTAATTACCGATCCATTAAAGGGGCCAGCCGCAAAA---GCTGGTATCAAGGCAGGTGAT 1172

Qy   1033 atcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtc 1092
Db   1173 GTTATTATTTCGGTAAATGGTGAGAAGATTAATGATGTCCGTGATCTAGCAAAGCGTATT 1232

Qy   1093 ggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatc 1152
Db   1233 GCAAATATGAGCCCAGGAGAAACAGTAACCTTAGGAGTTTGGAAATCTGGTAAAGAAGAG 1292

Qy   1153 acaatcaaagccaagctgg 1171
Db   1293 AATATTAAGGTTAAACTTG 1311


RESULT   7
PCT-US95-06211-7
; Sequence 7, Application PC/TUS9506211
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR DIAGNOSING
;    TITLE OF INVENTION:  ROCHALIMAEA HENSELAE AND ROCHALIMAEA QUINTANA INFECTION
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NEEDLE & ROSENBERG,P.C.
;      STREET:  127 Peachtree Street, Suite 1200
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06211
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/245,294
;      FILING DATE:  18 MAY 1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Spratt, Gwendolyn D.
;      REGISTRATION NUMBER:  36,016
;      REFERENCE/DOCKET NUMBER:  1414.6121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404/688-0770
;      TELEFAX:  404/688-9880
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1791 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  141..1649
PCT-US95-06211-7


  Query Match             11.1%;  Score 154.2;  DB 4;  Length 1791;
  Best Local Similarity   50.4%;  Pred. No. 2.1e-33;
  Matches  403;  Conservative    0;  Mismatches  393;  Indels    3;  Gaps    1;

Qy    373 ttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgtt 432
Db    516 TTTGGATCGGGTTTTTTTATCTCGTCTGATGGTTATATTGTGACCAATAATCATGTGATT 575

Qy    433 gccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactc 492
Db    576 TCTGATGGCACAAGTTACGCTGTTGTTCTTGATGACGGTACAGAACTGAATGCAAAACTC 635

Qy    493 atcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagcta 552
Db    636 ATTGGAACGGACCCACGAACTGATCTTGCAGTATTAAAAGTCAATGAAAAAAGAAAATTT 695

Qy    553 cccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatc 612
Db    696 TCGTACGTTGATTTTGGTGATGATTCAAAACTTCGTGTTGGTGATTGGGTTGTTGCTATT 755

Qy    613 ggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcaga 672
Db    756 GGTAATCCATTTGGTCTTGGTGGAACTGTGACAGCAGGTATCGTTTCAGCACGTGGACGT 815

Qy    673 agcctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggc 732
Db    816 GATATCGGTACCGGTGTTTATGATGATTTTATTCAGATTGATGCTGCAGTTAATCGAGGA 875

Qy    733 aattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaata 792
Db    876 AATTCTGGAGGTCCAACTTTTGATCTTAACGGAAAGGTTGTTGGAGTGAATACGGCAATT 935

Qy    793 tacagccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatg 852
Db    936 TTTTCTCCTTCTGGGGGCAACGTTGGGATTGCTTTCGCTATTCCGGCAGCAACAGCGAAC 995

Qy    853 aatgtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgatt 912
Db    996 GAGGTTGTGCAACAACTTATCGAAAAAGGTTTAGTTCAGCGTGGTTGGCTTGGGGTTCAG 1055

Qy    913 attcaggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgca 972
Db   1056 ATTCAGCCTGTAACAAAAGAAATTTCTGATTCAATTGGTTTGAAGGAGGCTAAAGGTGCG 1115

Qy    973 ttgattgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgac 1032
Db   1116 TTAATTACCGATCCATTAAAGGGGCCAGCCGCAAAA---GCTGGTATCAAGGCAGGTGAT 1172

Qy   1033 atcgtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtc 1092
Db   1173 GTTATTATTTCGGTAAATGGTGAGAAGATTAATGATGTCCGTGATCTAGCAAAGCGTATT 1232

Qy   1093 ggcgccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaagaaatc 1152
Db   1233 GCAAATATGAGCCCAGGAGAAACAGTAACCTTAGGAGTTTGGAAATCTGGTAAAGAAGAG 1292

Qy   1153 acaatcaaagccaagctgg 1171
Db   1293 AATATTAAGGTTAAACTTG 1311


RESULT   8
US-08-278-091-1
```

```
; Sequence 1, Application US/08278091
; Patent No. 5506139
;  GENERAL INFORMATION:
;    APPLICANT:  LOOSMORE, Sheena M
;    APPLICANT:  YANG, Yan-Ping
;    APPLICANT:  CHONG, Pele
;    APPLICANT:  OOMEN, Raymond P.
;    APPLICANT:  KLEIN, Michel H.
;    TITLE OF INVENTION:  Analog of Haemophilus Hin47 Protein with
;    TITLE OF INVENTION:  Reduced Protease Activity
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  Suite 701, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/278,091
;      FILING DATE:  21-JUL-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-371
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-278-091-1


  Query Match             10.0%;  Score 139.4;  DB 1;  Length 2894;
  Best Local Similarity   49.8%;  Pred. No. 3.1e-29;
  Matches  384;  Conservative    0;  Mismatches  381;  Indels    6;  Gaps    1;

Qy     382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
           | | | || |     |  || |||||  |  | |||||||  || ||  |||  |
Db    1108 GTCATTATTAATGCAAGCAAAGGCTATGTTTTAACCAATAATCATGTTATTGATGAAGCT 1167

Qy     442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
           |  |  || |  || |   |  | ||   ||| |||| ||  || ||| |  | |||
Db    1168 GATAAAATTACCGTGCAATTACAAGATGGGCGTGAATTTAAAGCAAAATTAGTGGGTAAA 1227

Qy     502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
           ||||  | ||| ||| | ||  |  |  |  | ||   | |     |  || | |   ||
Db    1228 GATGAACTATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287

Qy     562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
           ||| | |   || || | || ||       |||||  |     | ||| ||||    ||
Db    1288 AAATTTGCTGATTCCGACAAATTACGCGTAGGCGATTTCACTGTTGCAATCGGTAATCCA 1347

Qy     622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaagg------cagaagc 675
           || || || |   | |  |||||  | || || || || ||    ||       | | |
Db    1348 TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTCTGCATTGGGTCGTTCAACAGGT 1407

Qy     676 ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaat 735
             ||  |  |  ||   |  ||  || |||||||| |  ||  | || |  || |||
Db    1408 TCTGACAGTGGCACTTATGAAAACTATATTCAAACCGATGCAGCAGTAAACCGCGGTAAT 1467

Qy     736 tccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatac 795
           || || ||  || |  | ||||||| ||  |  |  | || || ||| |   |||
Db    1468 TCGGGTGGAGCGTTAGTAAACTTAAATGGCGAACTTATTGGAATTAATACCGCAATTATT 1527

Qy     796 agccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaat 855
             |  ||||| ||       || ||  ||||||| || || | | |    || |  |||
Db    1528 TCTCCAAGCGGTGGCAATGCAGGAATTGCCTTTGCGATTCCAAGTAATCAAGCAAGCAAT 1587

Qy     856 gtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattatt 915
             | |   ||||  |   | |    ||  |||| |  ||||||   || ||  | |    |
Db    1588 TTAGTGCAACAAATTTTAGAATTTGGTCAAGTGCGTCGCGGATTGCTTGGTATTAAAGGT 1647

Qy     916 caggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattg 975
             ||| |       | ||| ||  |  | ||   | |             ||||||||
Db    1648 GGCGAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAAGGCGCATTT 1707

Qy     976 attgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatc 1035
             |     || |  | ||        | || |||   || || ||  | ||||||| ||
Db    1708 GTAAGTGAAGTTTTACCGAAATCTGCTGCTGAAAAAGCAGGACTTAAAGCGGGCGATATT 1767

Qy    1036 gtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggc 1095
             ||       |  ||||   | ||||      || ||  ||  ||| |  |   |||
Db    1768 ATCACGGCGATGAACGGTCAAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy    1096 gccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaa 1146
             ||| |    | || |||||  | ||| |       |  ||  | ||| ||
Db    1828 ACCACTGGTGCAGGCAAAGAGATTAGCTTGACTTACTTACGTGATGGCAAA 1878


RESULT   9
US-08-483-859-1
; Sequence 1, Application US/08483859
; Patent No. 5656436
;  GENERAL INFORMATION:
;    APPLICANT:  LOOSMORE, Sheena M.
;    APPLICANT:  YANG, Yan-Ping
;    APPLICANT:  CHONG, Pele
;    APPLICANT:  OOMEN, Raymond P.
;    APPLICANT:  KLEIN, Michel H.
;    TITLE OF INVENTION:  Analog of Haemophilus Hin47 Protein with
;    TITLE OF INVENTION:  Reduced Protease Activity
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  Suite 701, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,859
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/296,149
;      FILING DATE:  26-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/278,091
;      FILING DATE:  21-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I.
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-495 MIS:vg
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2894 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-483-859-1


  Query Match             10.0%;  Score 139.4;  DB 1;  Length 2894;
  Best Local Similarity   49.8%;  Pred. No. 3.1e-29;
  Matches  384;  Conservative    0;  Mismatches  381;  Indels    6;  Gaps    1;

Qy     382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
           | | | || |     |  || ||||| |  | |||||||  || ||  |||  |
Db    1108 GTCATTATTAATGCAAGCAAAGGCTATGTTTTAACCAATAATCATGTTATTGATGAAGCT 1167

Qy     442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
           |  |  || |  || |   |  | ||   ||| |||| ||  || ||| |  | |||
Db    1168 GATAAAATTACCGTGCAATTACAAGATGGGCGTGAATTTAAAGCAAAATTAGTGGGTAAA 1227

Qy     502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
           ||||  | ||| ||| | ||  |  |  |  | ||   | |     |  || | |   ||
Db    1228 GATGAACTATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287

Qy     562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
           ||| | |   || || | || ||       ||||| |     | ||| |||||    ||
Db    1288 AAATTTGCTGATTCCGACAAATTACGCGTAGGCGATTTCACTGTTGCAATCGGTAATCCA 1347

Qy     622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaagg------cagaagc 675
           || || || |   | |  |||||  | || || || || || ||     ||       | | |
Db    1348 TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTCTGCATTGGGTCGTTCAACAGGT 1407

Qy     676 ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaat 735
             ||  |  |  ||   |  ||  || ||||||| |  ||  | || |  || |||
Db    1408 TCTGACAGTGGCACTTATGAAAACTATATTCAAACCGATGCAGCAGTAAACCGCGGTAAT 1467

Qy     736 tccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatac 795
           || || ||  || |  | ||||||| ||  |  |  || || ||| |   |||
Db    1468 TCGGGTGGAGCGTTAGTAAACTTAAATGGCGAACTTATTGGAATTAATACCGCAATTATT 1527

Qy     796 agccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaat 855
             |  ||||| ||       || ||  ||||||| || || | |    || |  |||
Db    1528 TCTCCAAGCGGTGGCAATGCAGGAATTGCCTTTGCGATTCCAAGTAATCAAGCAAGCAAT 1587

Qy     856 gtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattatt 915
           | |  ||||  |  | |    ||  |||| |  ||||||   || ||  | |     |
Db    1588 TTAGTGCAACAAATTTTAGAATTTGGTCAAGTGCGTCGCGGATTGCTTGGTATTAAAGGT 1647

Qy     916 caggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattg 975
             ||| |       | ||| ||  |  | ||   | |                ||||||||
Db    1648 GGCGAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAAGGCGCATTT 1707

Qy     976 attgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatc 1035
           |     || |  | ||        | || |||   || || ||  | |||||||| ||
Db    1708 GTAAGTGAAGTTTTACCGAAATCTGCTGCTGAAAAAGCAGGACTTAAAGCGGGCGATATT 1767

Qy    1036 gtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggc 1095
           ||       |  ||||   | ||||       || ||  ||  |||  |  |   |||
Db    1768 ATCACGGCGATGAACGGTCAAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy    1096 gccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaa 1146
           |||  |   | || |||||  | |||  |     |  ||  | |||  ||
Db    1828 ACCACTGGTGCAGGCAAAGAGATTAGCTTGACTTACTTACGTGATGGCAAA 1878


RESULT  10
US-08-472-173-1
; Sequence 1, Application US/08472173
; Patent No. 5665353
;  GENERAL INFORMATION:
;    APPLICANT:  LOOSMORE, Sheena M
;    APPLICANT:  YANG, Yan-Ping
;    APPLICANT:  CHONG, Pele
;    APPLICANT:  OOMEN, Raymond P.
;    APPLICANT:  KLEIN, Michel H.
;    TITLE OF INVENTION:  Analog of Haemophilus Hin47 Protein with
;    TITLE OF INVENTION:  Reduced Protease Activity
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  Suite 701, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,173
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/296,149
;      FILING DATE:  26-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/278,091
;      FILING DATE:  21-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I.
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-493 MIS:vg
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-472-173-1


  Query Match             10.0%;  Score 139.4;  DB 1;  Length 2894;
  Best Local Similarity   49.8%;  Pred. No. 3.1e-29;
  Matches  384;  Conservative    0;  Mismatches  381;  Indels    6;  Gaps    1;

Qy     382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
           | | | || |     |  || ||||| |  | |||||||  || ||  |||  |
Db    1108 GTCATTATTAATGCAAGCAAAGGCTATGTTTTAACCAATAATCATGTTATTGATGAAGCT 1167

Qy     442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
           |  |  || |  || |   |  | ||   ||| |||| ||  || ||| |  | |||
Db    1168 GATAAAATTACCGTGCAATTACAAGATGGGCGTGAATTTAAAGCAAAATTAGTGGGTAAA 1227

Qy     502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
           ||||  | ||| ||| | ||  |  |  |  | ||   | |     |  || | |   ||
Db    1228 GATGAACTATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287

Qy     562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
           ||| | |   || || | || ||       ||||| |     | ||| |||||    ||
Db    1288 AAATTTGCTGATTCCGACAAATTACGCGTAGGCGATTTCACTGTTGCAATCGGTAATCCA 1347

Qy     622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaagg------cagaagc 675
           || || || |   | |  |||||  | || || || || || ||     ||       | | |
Db    1348 TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTCTGCATTGGGTCGTTCAACAGGT 1407

Qy     676 ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaat 735
             ||  |  |  ||   |  ||  || ||||||| |  ||  | || |  || |||
Db    1408 TCTGACAGTGGCACTTATGAAAACTATATTCAAACCGATGCAGCAGTAAACCGCGGTAAT 1467

Qy     736 tccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatac 795
           || || ||  || |  | ||||||| ||  |  |  || || ||| |   |||
Db    1468 TCGGGTGGAGCGTTAGTAAACTTAAATGGCGAACTTATTGGAATTAATACCGCAATTATT 1527

Qy     796 agccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaat 855
```

```
Db   1528 TCTCCAAGCGGTGGCAATGCAGGAATTGCCTTTGCGATTCCAAGTAATCAAGCAAGCAAT 1587

Qy    856 gtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattatt 915
Db   1588 TTAGTGCAACAAATTTTAGAATTTGGTCAAGTGCGTCGCGGATTGCTTGGTATTAAAGGT 1647

Qy    916 caggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattg 975
Db   1648 GGCGAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAAGGCGCATTT 1707

Qy    976 attgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatc 1035
Db   1708 GTAAGTGAAGTTTTACCGAAATCTGCTGCTGAAAAAGCAGGACTTAAAGCGGGCGATATT 1767

Qy   1036 gtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggc 1095
Db   1768 ATCACGGCGATGAACGGTCAAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy   1096 gccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaa 1146
Db   1828 ACCACTGGTGCAGGCAAAGAGATTAGCTTGACTTACTTACGTGATGGCAAA 1878


RESULT 11
US-08-487-167-1
; Sequence 1, Application US/08487167
; Patent No. 5869302
;  GENERAL INFORMATION:
;    APPLICANT:  LOOSMORE, Sheena M.
;    APPLICANT:  YANG, Yan-Ping
;    APPLICANT:  CHONG, Pele
;    APPLICANT:  OOMEN, Raymond P.
;    APPLICANT:  KLEIN, Michel H.
;    TITLE OF INVENTION:  Analog of Haemophilus Hin47 Protein with
;    TITLE OF INVENTION:  Reduced Protease Activity
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  Suite 701, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,167
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/296,149
;      FILING DATE:  26-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/278,091
;      FILING DATE:  21-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I.
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-508 MIS:vg
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-487-167-1

  Query Match             10.0%;  Score 139.4;  DB 2;  Length 2894;
  Best Local Similarity   49.8%;  Pred. No. 3.1e-29;
  Matches  384;  Conservative    0;  Mismatches  381;  Indels    6;  Gaps    1;

Qy    382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
Db   1108 GTCATTATTAATGCAAGCAAAGGCTATGTTTTAACCAATAATCATGTTATTGATGAAGCT 1167

Qy    442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
Db   1168 GATAAAATTACCGTGCAATTACAAGATGGGCGTGAATTTAAAGCAAAATTAGTGGGTAAA 1227

Qy    502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
Db   1228 GATGAACTATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287

Qy    562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
Db   1288 AAATTTGCTGATTCCGACAAATTACGCGTAGGCGATTTCACTGTTGCAATCGGTAATCCA 1347

Qy    622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaagg------cagaagc 675
Db   1348 TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTCTGCATTGGGTCGTTCAACAGGT 1407

Qy    676 ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaat 735
Db   1408 TCTGACAGTGGCACTTATGAAAACTATATTCAAACCGATGCAGCAGTAAACCGCGGTAAT 1467

Qy    736 tccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatac 795
Db   1468 TCGGGTGGAGCGTTAGTAAACTTAAATGGCGAACTTATTGGAATTAATACCGCAATTATT 1527

Qy    796 agccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaat 855
Db   1528 TCTCCAAGCGGTGGCAATGCAGGAATTGCCTTTGCGATTCCAAGTAATCAAGCAAGCAAT 1587

Qy    856 gtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattatt 915
Db   1588 TTAGTGCAACAAATTTTAGAATTTGGTCAAGTGCGTCGCGGATTGCTTGGTATTAAAGGT 1647

Qy    916 caggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattg 975
Db   1648 GGCGAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAAGGCGCATTT 1707

Qy    976 attgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatc 1035
Db   1708 GTAAGTGAAGTTTTACCGAAATCTGCTGCTGAAAAAGCAGGACTTAAAGCGGGCGATATT 1767

Qy   1036 gtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggc 1095
Db   1768 ATCACGGCGATGAACGGTCAAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy   1096 gccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaa 1146
Db   1828 ACCACTGGTGCAGGCAAAGAGATTAGCTTGACTTACTTACGTGATGGCAAA 1878


RESULT 12
US-08-482-816-1
; Sequence 1, Application US/08482816
; Patent No. 5935573
;  GENERAL INFORMATION:
;    APPLICANT:  LOOSMORE, Sheena M
;    APPLICANT:  YANG, Yan-Ping
;    APPLICANT:  CHONG, Pele
;    APPLICANT:  OOMEN, Raymond P.
;    APPLICANT:  KLEIN, Michel H.
;    TITLE OF INVENTION:  Analog of Haemophilus Hin47 Protein with Reduced Protease A
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
```

```
;      STREET:  Suite 701, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,816
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/296,149
;      FILING DATE:  26-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/278,091
;      FILING DATE:  21-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-494 MIS:vg
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-482-816-1


  Query Match           10.0%;  Score 139.4;  DB 2;  Length 2894;
  Best Local Similarity   49.8%;  Pred. No. 3.1e-29;
  Matches  384;  Conservative    0;  Mismatches  381;  Indels    6;  Gaps    1;

Qy     382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
          | | | || |     |  || |||||  |  | |||||||  || ||  |||  |
Db    1108 GTCATTATTAATGCAAGCAAAGGCTATGTTTTAACCAATAATCATGTTATTGATGAAGCT 1167

Qy     442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
           |  |  || |  || |    |  | ||    ||| |||| ||  || ||| |  | |||
Db    1168 GATAAAATTACCGTGCAATTACAAGATGGGCGTGAATTTAAAGCAAAATTAGTGGGTAAA 1227

Qy     502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
           ||||  | ||| ||| | ||  |  |  |  | ||   | |     |  || | |  ||
Db    1228 GATGAACTATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287

Qy     562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
           ||| | |   || || | || ||       |||| |     | ||| |||||    ||
Db    1288 AAATTTGCTGATTCCGACAAATTACGCGTAGGCGATTTCACTGTTGCAATCGGTAATCCA 1347

Qy     622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaagg------cagaagc 675
           || || || |   | |  |||||  | || || || || || ||    ||       | | |
Db    1348 TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTCTGCATTGGGTCGTTCAACAGGT 1407

Qy     676 ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaat 735
                ||  |  |  ||   |  ||  || |||||||| |  ||  | || |  || |||
Db    1408 TCTGACAGTGGCACTTATGAAAACTATATTCAAACCGATGCAGCAGTAAACCGCGGTAAT 1467

Qy     736 tccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatac 795
           || || ||  || |  | |||||||| ||  |  |  | || || ||| |   |||
Db    1468 TCGGGTGGAGCGTTAGTAAACTTAAATGGCGAACTTATTGGAATTAATACCGCAATTATT 1527

Qy     796 agccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaat 855
             |  ||||| ||       || ||  |||||||| || || | | |    || |  |||
Db    1528 TCTCCAAGCGGTGGCAATGCAGGAATTGCCTTTGCGATTCCAAGTAATCAAGCAAGCAAT 1587

Qy     856 gtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattatt 915
```

```
           | |   ||||  |   | |    ||  |||| |  ||||||   || ||  | |    |
Db    1588 TTAGTGCAACAAATTTTAGAATTTGGTCAAGTGCGTCGCGGATTGCTTGGTATTAAAGGT 1647

Qy     916 caggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattg 975
             ||| |       | ||| ||  |  | ||   | |              |||||||
Db    1648 GGCGAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAAGGCGCATTT 1707

Qy     976 attgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatc 1035
           |     || |  | ||         | || |||   || || ||  | ||||||| ||
Db    1708 GTAAGTGAAGTTTTACCGAAATCTGCTGCTGAAAAAGCAGGACTTAAAGCGGGCGATATT 1767

Qy    1036 gtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggc 1095
           ||       | ||||  | ||||      || ||  ||  |||  |  |   |||
Db    1768 ATCACGGCGATGAACGGTCAAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy    1096 gccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaa 1146
           |||  |    | || |||||  | |||  |       |  ||  | ||| ||
Db    1828 ACCACTGGTGCAGGCAAAGAGATTAGCTTGACTTACTTACGTGATGGCAAA 1878


RESULT  13
US-08-296-149-1
; Sequence 1, Application US/08296149
; Patent No. 5939297
;  GENERAL INFORMATION:
;    APPLICANT:  LOOSMORE, Sheena M
;    APPLICANT:  YANG, Yan-Ping
;    APPLICANT:  CHONG, Pele
;    APPLICANT:  OOMEN, Raymond P.
;    APPLICANT:  KLEIN, Michel H.
;    TITLE OF INVENTION:  Analog of Haemophilus Hin47 Protein with
;    TITLE OF INVENTION:  Reduced Protease Activity
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  Suite 701, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/296,149
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-390
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-296-149-1


  Query Match           10.0%;  Score 139.4;  DB 2;  Length 2894;
  Best Local Similarity   49.8%;  Pred. No. 3.1e-29;
  Matches  384;  Conservative    0;  Mismatches  381;  Indels    6;  Gaps    1;

Qy     382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
          | | | || |     |  || |||||  |  | |||||||  || ||  |||  |
Db    1108 GTCATTATTAATGCAAGCAAAGGCTATGTTTTAACCAATAATCATGTTATTGATGAAGCT 1167
```

```
Qy     442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
           |  |  || |  || |    |  | ||   ||| |||| ||  || ||| |  | |||
Db    1168 GATAAAATTACCGTGCAATTACAAGATGGGCGTGAATTTAAAGCAAAATTAGTGGGTAAA 1227

Qy     502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
           ||||  | ||| ||| | ||  |  |  |  | ||   | |     |  || | |   ||
Db    1228 GATGAACTATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287

Qy     562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
           ||| | |   || || | || ||       ||||| |     | ||| |||||    ||
Db    1288 AAATTTGCTGATTCCGACAAATTACGCGTAGGCGATTTCACTGTTGCAATCGGTAATCCA 1347

Qy     622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaagg------cagaagc 675
           || || || |   | |  ||||| | || || || || ||    ||       | | |
Db    1348 TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTCTGCATTGGGTCGTTCAACAGGT 1407

Qy     676 ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaat 735
             ||  |  |  ||   |  ||  || |||||||| |  ||  | || |  || |||
Db    1408 TCTGACAGTGGCACTTATGAAAACTATATTCAAACCGATGCAGCAGTAAACCGCGGTAAT 1467

Qy     736 tccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatac 795
           || || ||  || |  | |||||||| ||  |  |  | || || ||| |   |||
Db    1468 TCGGGTGGAGCGTTAGTAAACTTAAATGGCGAACTTATTGGAATTAATACCGCAATTATT 1527

Qy     796 agccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaat 855
              |  ||||| ||       || ||  ||||||| || || | | |    || |  |||
Db    1528 TCTCCAAGCGGTGGCAATGCAGGAATTGCCTTTGCGATTCCAAGTAATCAAGCAAGCAAT 1587

Qy     856 gtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattatt 915
           | |   ||||  |   | |    ||  |||| |  ||||||   || ||  | |    |
Db    1588 TTAGTGCAACAAATTTTAGAATTTGGTCAAGTGCGTCGCGGATTGCTTGGTATTAAAGGT 1647

Qy     916 caggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattg 975
             ||| |       | ||| ||  |  | ||  | |                |||||||
Db    1648 GGCGAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAAGGCGCATTT 1707

Qy     976 attgccaaaatccttcccggcagcccccgcagaacgtgccggcctgcaggcgggcgacatc 1035
             |     || |  | ||        | || |||   || || ||  | ||||||| ||
Db    1708 GTAAGTGAAGTTTTACCGAAATCTGCTGCTGAAAAAGCAGGACTTAAAGCGGGCGATATT 1767

Qy    1036 gtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggc 1095
           ||       |  ||||   | ||||      || ||  ||  |||  |  |   |||
Db    1768 ATCACGGCGATGAACGGTCAAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy    1096 gccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaa 1146
           |||  |    | || |||||  | |||  |       |  ||  | ||| ||
Db    1828 ACCACTGGTGCAGGCAAAGAGATTAGCTTGACTTACTTACGTGATGGCAAA 1878
```

```
RESULT  14
US-08-801-499-1
; Sequence 1, Application US/08801499
; Patent No. 5962430
;  GENERAL INFORMATION:
;    APPLICANT:  LOOSMORE, Sheena M
;    APPLICANT:  YANG, Yan-Ping
;    APPLICANT:  CHONG, Pele
;    APPLICANT:  OOMEN, Raymond P.
;    APPLICANT:  KLEIN, Michel H.
;    TITLE OF INVENTION:  Analog of Haemophilus Hin47 Protein with Reduced Protease Act
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  6th Floor, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/801,499
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/482,816
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/296,149
;      FILING DATE:  26-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/278,091
;      FILING DATE:  21-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-671 MIS:jb
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-801-499-1


  Query Match             10.0%;  Score 139.4;  DB 2;  Length 2894;
  Best Local Similarity   49.8%;  Pred. No. 3.1e-29;
  Matches  384;  Conservative    0;  Mismatches  381;  Indels    6;  Gaps    1;

Qy     382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
           | | | || |     |  || |||||  |  | |||||||  || ||  |||  |
Db    1108 GTCATTATTAATGCAAGCAAAGGCTATGTTTTAACCAATAATCATGTTATTGATGAAGCT 1167

Qy     442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
           |  |  || |  || |    |  | ||   ||| |||| ||  || ||| |  | |||     '
Db    1168 GATAAAATTACCGTGCAATTACAAGATGGGCGTGAATTTAAAGCAAAATTAGTGGGTAAA 1227

Qy     502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
           ||||  | ||| ||| | ||  |  |  |  | ||   | |     |  || | |   ||
Db    1228 GATGAACTATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287

Qy     562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
           ||| | |   || || | || ||       ||||| |     | ||| |||||    ||
Db    1288 AAATTTGCTGATTCCGACAAATTACGCGTAGGCGATTTCACTGTTGCAATCGGTAATCCA 1347

Qy     622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaagg------cagaagc 675
           || || || |   | |  ||||| | || || || || ||    ||       | | |
Db    1348 TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTCTGCATTGGGTCGTTCAACAGGT 1407

Qy     676 ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaat 735
             ||  |  |  ||   |  ||  || |||||||| |  ||  | || |  || |||
Db    1408 TCTGACAGTGGCACTTATGAAAACTATATTCAAACCGATGCAGCAGTAAACCGCGGTAAT 1467

Qy     736 tccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatac 795
           || || ||  || |  | |||||||| ||  |  |  | || || ||| |   |||
Db    1468 TCGGGTGGAGCGTTAGTAAACTTAAATGGCGAACTTATTGGAATTAATACCGCAATTATT 1527

Qy     796 agccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaat 855
              |  ||||| ||       || ||  ||||||| || || | | |    || |  |||
Db    1528 TCTCCAAGCGGTGGCAATGCAGGAATTGCCTTTGCGATTCCAAGTAATCAAGCAAGCAAT 1587

Qy     856 gtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattatt 915
           | |   ||||  |   | |    ||  |||| |  ||||||   || ||  | |    |
Db    1588 TTAGTGCAACAAATTTTAGAATTTGGTCAAGTGCGTCGCGGATTGCTTGGTATTAAAGGT 1647

Qy     916 caggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattg 975
             ||| |       | ||| ||  |  | ||  | |                |||||||
Db    1648 GGCGAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAAGGCGCATTT 1707
```

```
Qy    976 attgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatc 1035
          |     || |  | ||        | || |||   || || ||  | ||||||| ||
Db   1708 GTAAGTGAAGTTTTACCGAAATCTGCTGCTGAAAAAGCAGGACTTAAAGCGGGCGATATT 1767

Qy   1036 gtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggc 1095
          ||       |  ||||   | ||||      || ||  ||  |||  |  |   |||
Db   1768 ATCACGGCGATGAACGGTCAAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy   1096 gccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaa 1146
          ||| |    | || ||||  | ||| |       |  ||  | ||| ||
Db   1828 ACCACTGGTGCAGGCAAAGAGATTAGCTTGACTTACTTACGTGATGGCAAA 1878


RESULT  15
US-08-615-271-1
; Sequence 1, Application US/08615271
; Patent No. 5981503
;  GENERAL INFORMATION:
;    APPLICANT:  LOOSMORE, Sheena M.
;    APPLICANT:  YANG, Yan-Ping
;    APPLICANT:  CHONG, Pele
;    APPLICANT:  OOMEN, Raymond P.
;    APPLICANT:  KLEIN, Michel H.
;    TITLE OF INVENTION:  ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED
;    TITLE OF INVENTION:  PROTEASE ACTIVITY
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  6th Floor, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/615,271
;      FILING DATE:  20-JUN-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-580
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2894 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-615-271-1


  Query Match             10.0%;  Score 139.4;  DB 2;  Length 2894;
  Best Local Similarity   49.8%;  Pred. No. 3.1e-29;
  Matches  384;  Conservative    0;  Mismatches  381;  Indels    6;  Gaps    1;

Qy    382 ggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgccggtatg 441
          | | | || |     |  || |||||  |  | |||||||  || ||  |||  |
Db   1108 GTCATTATTAATGCAAGCAAAGGCTATGTTTTAACCAATAATCATGTTATTGATGAAGCT 1167

Qy    442 ggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatcggttcg 501
          |  |  || |  || |   |  | ||   ||| |||| ||  || ||| |  | |||
Db   1168 GATAAAATTACCGTGCAATTACAAGATGGGCGTGAATTTAAAGCAAAATTAGTGGGTAAA 1227

Qy    502 gatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtc 561
          ||||  | ||| ||| | ||  |  |  |  | ||   | |     |  || | |   ||

Db   1228 GATGAACTATCAGATATTGCATTAGTACAGCTTGAAAAACCAAGTAATTTAACAGAAATC 1287

Qy    562 aaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgccc 621
          ||| | |   || || | || ||        ||||| |     | ||| |||||    ||
Db   1288 AAATTTGCTGATTCCGACAAATTACGCGTAGGCGATTTCACTGTTGCAATCGGTAATCCA 1347

Qy    622 ttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaagg------cagaagc 675
          || || || |   | |  |||||  | || || || || ||    ||          | | |
Db   1348 TTTGGTTTAGGTCAAACTGTGACATCAGGTATTGTTTCTGCATTGGGTCGTTCAACAGGT 1407

Qy    676 ctgcccaacgaaagctacacaccttcatccaaaccgacgttgccatcaatccgggcaat 735
              ||  |  |  ||   |  ||  || ||||||||  |  ||  | || |  || |||
Db   1408 TCTGACAGTGGCACTTATGAAAACTATATTCAAACCGATGCAGCAGTAAACCGCCGTAAT 1467

Qy    736 tccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaattcgcaaatatac 795
          || || ||  || |  | ||||||||  ||  |  |  | || || ||| |   |||
Db   1468 TCGGGTGGAGCGTTAGTAAACTTAAATGGCGAACTTATTGGAATTAATACCGCAATTATT 1527

Qy    796 agccgcagcggcggattcatgggcatctcctttgccatcccgattgacgttgccatgaat 855
            |  ||||| ||       || ||  |||||||  || || | | |    || |  |||
Db   1528 TCTCCAAGCGGTGGCAATGCAGGAATTGCCTTTGCGATTCCAAGTAATCAAGCAAGCAAT 1587

Qy    856 gtcgccgaacagctgaaaaacaccggcaaagtccaacgcggacaactgggcgtgattatt 915
          | |   |||| |   | |    ||  |||| |  ||||||   || ||  | |    |
Db   1588 TTAGTGCAACAAATTTTAGAATTTGGTCAAGTGCGTCGCGGATTGCTTGGTATTAAAGGT 1647

Qy    916 caggaagtatcctacggtttggcacagtcgttcggtctggataaagccagcggcgcattg 975
          ||| |       | ||| ||  |  | ||   | |                |||||||
Db   1648 GGCGAACTCAATGCTGATTTAGCCAAAGCCTTTAATGTAAGCGCGCAACAAGGCGCATTT 1707

Qy    976 attgccaaaatccttcccggcagccccgcagaacgtgccggcctgcaggcgggcgacatc 1035
          |     || |  | ||        | || |||   || || ||  | ||||||| ||
Db   1708 GTAAGTGAAGTTTTACCGAAATCTGCTGCTGAAAAAGCAGGACTTAAAGCGGGCGATATT 1767

Qy   1036 gtcctcagcctcgacggcggagaaatacgttcttccggcgaccttcccgtcatggtcggc 1095
          ||       |  ||||   | ||||      || ||  ||  |||  |  |   |||
Db   1768 ATCACGGCGATGAACGGTCAAAAAATCTCAAGTTTCGCTGAAATTCGTGCAAAAATCGCA 1827

Qy   1096 gccattacgccgggaaaagaagtcagcctcggcgtatggcgcaaaggcgaa 1146
          ||| |    | || ||||  | ||| |       |  ||  | ||| ||
Db   1828 ACCACTGGTGCAGGCAAAGAGATTAGCTTGACTTACTTACGTGATGGCAAA 1878


Search completed: March 22, 2001, 04:05:10
Job time: 4435 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: March 22, 2001, 02:20:35 ; Search time 899.71 Seconds
(without alignments)
10865.107 Million cell updates/sec

Title: US-09-388-090-3
Perfect score: 1395
Sequence: 1 gtgttcaaaaaataccaata.....tcagggcaagtccccgtcaa 1395

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*
44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estpl1:*
70: em_estpl2:*
71: em_estpl3:*
72: em_estpl4:*
73: em_estpl5:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estpl6:*
118: em_estpl7:*
119: em_estpl8:*
120: em_estro14:*
121: em_estro15:*
122: em_estro16:*
123: em_estro17:*
124: em_estro18:*
125: em_estro19:*
126: gb_est58:*
127: gb_est59:*
128: gb_est60:*
129: gb_est61:*
130: gb_est62:*
131: gb_est63:*
132: gb_est64:*
133: gb_est65:*
134: gb_est66:*
135: gb_est75:*
136: gb_est76:*
137: gb_est77:*
138: gb_est78:*
139: gb_est79:*
140: gb_est80:*
141: gb_est81:*
142: gb_est82:*
143: gb_est83:*
144: gb_est84:*
145: gb_est85:*
146: gb_est86:*
147: gb_est87:*
148: gb_gss1:*
149: gb_gss2:*
150: gb_gss3:*
151: gb_gss4:*
152: em_gss1:*
153: em_gss2:*
154: em_gss3:*
155: em_gss4:*
156: gb_gss5:*
157: gb_gss6:*
158: gb_gss7:*
159: gb_gss8:*
160: gb_gss9:*
161: gb_gss10:*
162: gb_gss11:*
163: gb_gss12:*
164: gb_gss13:*
165: gb_gss14:*
166: gb_gss15:*
167: gb_gss16:*
168: gb_gss17:*
169: gb_gss18:*
170: gb_gss19:*
171: gb_gss20:*
172: gb_gss21:*
173: gb_gss22:*
174: gb_gss23:*
175: gb_gss24:*
176: em_gss5:*
177: em_gss6:*
178: em_gss7:*
179: em_gss8:*
180: em_gss9:*
181: em_gss10:*
182: em_gss11:*
183: em_gss12:*
184: em_gss13:*
185: em_gss14:*
186: em_gss15:*
187: em_gss16:*
188: em_gss17:*
189: em_gss18:*
190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	100.6	7.2	490	8	AA479844	AA479844 zu43h07.r
c	2	98.6	7.1	479	91	AW491392	AW491392 U1-M-BH3-
	3	97.4	7.0	489	2	AA110074	AA110074 mo51b12.r
	4	97.4	7.0	660	94	AW742283	AW742283 up55a10.y
	5	95.4	6.8	456	16	AI159029	AI159029 vz83a08.r
	6	94.6	6.8	521	93	AW654015	AW654015 103226 MA
	7	92.2	6.6	474	146	W17262	W17262 zb17a08.r1
	8	92.2	6.6	485	144	R52327	R52327 yg75a09.r1
	9	91	6.5	504	93	AW660844	AW660844 99933 MAR
	10	89.6	6.4	350	140	F07134	F07134 HSC1WG081 n
	11	89	6.4	529	37	AV614726	AV614726 AV614726
	12	87.8	6.3	559	37	AV599976	AV599976 AV599976
c	13	86.2	6.2	337	26	AI903651	AI903651 QV-BT032-
	14	84.2	6.0	518	144	R22703	R22703 yg08e06.r1
	15	81.2	5.8	288	146	T80106	T80106 yd04h09.r1
	16	81.2	5.8	346	139	C18770	C18770 C18770 Huma
	17	80.2	5.7	500	38	AV667427	AV667427 AV667427
	18	78.8	5.6	375	10	AA644825	AA644825 vs83a09.r
	19	78	5.6	689	14	AF179474	AF179474 AF179474
	20	77.8	5.6	623	27	AI982113	AI982113 pat.pk007
	21	77.2	5.5	280	89	AW326487	AW326487 19115 MAR
	22	75.8	5.4	573	105	BE236269	BE236269 143992 MA
	23	75.8	5.4	574	105	BE236299	BE236299 144028 MA
c	24	75	5.4	402	37	AV614725	AV614725 AV614725
c	25	72.4	5.2	509	90	AW462478	AW462478 BP230009B
	26	69.6	5.0	292	6	AA349615	AA349615 EST56620
	27	69.4	5.0	485	3	AA186646	AA186646 zp64g02.r
	28	66.6	4.8	281	140	F08513	F08513 HSCZWA091 n
	29	66.2	4.7	483	107	BE335671	BE335671 us87h06.y
	30	65.6	4.7	508	141	H42958	H42958 yo68e01.r1
	31	65	4.7	496	38	AW013197	AW013197 Sp66f Win
	32	64.2	4.6	591	37	AV593667	AV593667 AV593667
	33	64	4.6	544	134	BE026352	BE026352 db37c11.y
	34	62.6	4.5	669	8	AA545554	AA545554 vj90d11.s
	35	61.8	4.4	511	1	AA028600	AA028600 mi14a10.r
	36	59.2	4.2	442	141	H16043	H16043 yl27g02.r1
	37	57.4	4.1	504	13	AA913980	AA913980 vy91c06.r
	38	57	4.1	275	5	AA340884	AA340884 EST46223
	39	56.8	4.1	502	135	BE754831	BE754831 208470 MA
	40	55.8	4.0	383	90	AW462230	AW462230 BP230008B
	41	55.6	4.0	298	4	AA254878	AA254878 mz75h12.r
	42	54.4	3.9	341	136	BE843509	BE843509 CM3-TN006
	43	53.2	3.8	617	134	BE055938	BE055938 945020D03
	44	52.2	3.7	427	93	AW632119	AW632119 91674 MAR
	45	52.2	3.7	453	90	AW446594	AW446594 86442 MAR

ALIGNMENTS

RESULT 1
AA479844
LOCUS AA479844 490 bp mRNA EST 09-NOV-1997
DEFINITION zu43h07.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:740797 5' similar to SW:HTRA_SALTY P26982 PROTEASE DO PRECURSOR ;, mRNA sequence.
ACCESSION AA479844
VERSION AA479844.1 GI:2205730
KEYWORDS EST.

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 490)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2408   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 417.
FEATURES             Location/Qualifiers
     source          1. .490
                     /organism="Homo sapiens"
                     /db_xref="GDB:5940721"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:740797"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT      138 a    132 c    127 g     93 t
ORIGIN
```

Query Match 7.2%; Score 100.6; DB 8; Length 490;
Best Local Similarity 55.7%; Pred. No. 6.7e-18;
Matches 221; Conservative 0; Mismatches 164; Indels 12; Gaps 1;

```
Qy    398 aaaacggctacatcctgaccaatacccacgtcgttgccggtatgggcagtatcaaagtcc 457
          || | ||    ||| |||| ||| ||||||| ||  ||    | |  | |  |||||||
Db     11 AAGATGGACTGATCGTGACAAATGCCCACGTGGTGACCAACAAGCACCGGGTCAAAGTTG 70

Qy    458 tgctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatccgatg 517
           ||| ||  ||     |   ||    |||||| |||  | |  |||||   || | ||
Db     71 AGCTGAAGAACGGTGCCACTTACGAAGCCAAAATCAAGGATGTGGATGAGAAAGCAGACA 130

Qy    518 tcgcccttctgaaaatcgacgcaacggaagagctacccgtcgtcaaaatcggcaatccca 577
          |||| ||  | ||||| |||     ||   |||| || ||| |     | |||    ||
Db    131 TCGCACTCATCAAAATTGACCACCAGGGCAAGCTGCCTGTCCTGCTGCTTGGCCGCTCCT 190

Qy    578 aaaatttgaaaccgggcgaatgggtcgctgccatcggcgcgcccttcggctttgacaaca 637
          | |  ||   ||||| || |  || |  ||||||||    || ||   | || | ||||
Db    191 CAGAGCTGCGGCCGGGAGAGTTCGTGGTCGCCATCGGAAGCCCGTTTTCCCTTCAAAACA 250

Qy    638 gcgtgaccgccggcatcgtgtccgccaaaggcagaagc----------ctgcccaacg 685
           || ||| |||| ||||||  | |||       || ||             || |  |
Db    251 CAGTCACCACCGGGATCGTGAGCACCACCCAGCGAGGCGACAAAGAGCTGGGGCTCCGCA 310

Qy    686 aaagctacacacccttcatccaaaccgacgttgccatcaatccgggcaattccggcggcc 745
          |     |||    || |||||| |||||||    ||||||    || || || || ||||
Db    311 ACTCAGACATGGACTACATCCAGACCGACGCCATCATCAACTATGGAAACTCGGGAGGCC 370

Qy    746 cgctgttcaacttaaaaggacaggtcgtcggcatcaa 782
          || |  | ||| |  | ||  | ||  | || || ||
Db    371 CGTTAGTAAACCTGGACGGTGAAGTGATTGGAATTAA 407
```

```
RESULT   2
AW491392/c
LOCUS       AW491392      479 bp    mRNA            EST       24-FEB-2000
DEFINITION  UI-M-BH3-atr-e-02-0-UI.s1 NIH_BMAP_M_S4 Mus musculus cDNA clone
            UI-M-BH3-atr-e-02-0-UI 3', mRNA sequence.
ACCESSION   AW491392
VERSION     AW491392.1  GI:7061638
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 479)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to identify it as a clone from the
            normalized corpus striatum library cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
            clones from RESEARCH GENETICS. It should be noted that Bento Soares
            is generating a small number of additional specialized
            non-redundant arrays of BMAP cDNAs whose availability will be
            considered under appropriate and limited collaborative arrangements
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .479
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH3-atr-e-02-0-UI"
                     /clone_lib="NIH_BMAP_M_S4"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_M_S4 library is a subtracted library of a series,
                     ultimately derived from a mixture of individually tagged
                     normalized libraries from ten regions of the mouse brain
                     (cerebellum, brain stems, olfactory bulbs, hypothalamus,
                     cortex, amygdala, basal ganglia, pineal gland, striatum,
                     hipoccampus) after a series of subtractions to reduce the
                     representation of cDNAs from which ESTs had already been
                     generated. The following serially subtracted libraries
                     were generated in this process: NIH_BMAP_M_S4,
                     NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
                     NIH_BMAP_M_S2, NIH_BMAP_M_S1.  The subtracted library
                     (NIH_BMAP_M_S4) was constructed as follows:  PCRamplified
                     cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
                     NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
                     was used as  a driver in a hybridization with a pool of
                     the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
                     libraries in the form of  single-stranded circles.  The
                     remaining single-stranded circles  (subtracted library)
                     was purified by hydroxyapatite column chromatography,
```

converted to double-stranded circles and electroporated into DH10B bacteria (LifeTechnologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_LIB=NIH_BMAP_M_S4
TAG_TISSUE=corpus-striatum
TAG_SEQ=ACGGC"
BASE COUNT 99 a 126 c 111 g 143 t
ORIGIN

Query Match 7.1%; Score 98.6; DB 91; Length 479;
Best Local Similarity 54.4%; Pred. No. 2.5e-17;
Matches 228; Conservative 0; Mismatches 179; Indels 12; Gaps 1;

```
Qy    376 ggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgcc 435
          || || || |||||| |     |  | ||    ||  |||| ||| | ||||| ||  ||
Db    475 GGGTCAGGATTCATCGTATCGGAGGATGGACTGATTGTGACAAATGCTCACGTGGTCACC 416

Qy    436 ggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatc 495
            |  | |  |||| ||   ||| ||  |          |||   |||||| |||
Db    415 AACAAAAACCGGGTCAAGGTTGAGCTGAAGAATGGAGCTACCTATGAAGCCAAAATCAAG 356

Qy    496 ggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctaccc 555
          | |  |||||    |  | ||  | || ||| | || || |||     || | |||| ||
Db    355 GATGTGGATGAAAAGGCGGACATTGCGCTTATCAAGATTGACCACCAGGGAAAGCTGCCA 296

Qy    556 gtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggc 615
          ||| |     |||||    ||  | |  ||| ||| || ||||  || | ||||||  ||
Db    295 GTCCTGCTGCTCGGCCGCTCCTCAGAGCTGAGACCTGGAGAATTTGTAGTTGCCATTGGA 236

Qy    616 gcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcagaagc 675
            |||||    || | ||||  || ||| | || |||||   | |||     || ||
Db    235 AGCCCCTTTTCTCTTCAAAACACAGTCACCACTGGGATCGTCAGCACCACCCAGCGAGGC 176

Qy    676 ------------ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatc 723
                       |    |  | | |    || ||| || || |||| |  ||||
Db    175 GGCAAAGAGCTGGGACTTCGGAACTCCGATATGGACTACATTCAGACAGACGCTATCATC 116

Qy    724 aatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaa 782
          |||   || |||||||| ||||||  |  | ||| |  | ||  ||||  | || || ||
Db    115 AATTATGGAAATTCCGGAGGCCCGTTAGTAAACCTGGATGGCGAGGTGATTGGGATTAA 57
```

RESULT 3
AA110074
LOCUS AA110074 489 bp mRNA EST 06-NOV-1996
DEFINITION mo51b12.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus cDNA clone IMAGE:557087 5' similar to SW:HTRA_SALTY P26982 PROTEASE DO PRECURSOR ;, mRNA sequence.
ACCESSION AA110074
VERSION AA110074.1 GI:1661822
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 489)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:337879
Seq primer: -28M13 rev1 from Amersham
High quality sequence stop: 446.
FEATURES Location/Qualifiers
source 1. .489
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:557087"
/clone_lib="Life Tech mouse embryo 10 5dpc 10665016"
/tissue_type="embryo"
/dev_stage="10.5dpc embryos"
/lab_host="DH10B"
/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. 10.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT 130 a 116 c 140 g 103 t
ORIGIN

Query Match 7.0%; Score 97.4; DB 2; Length 489;
Best Local Similarity 53.8%; Pred. No. 5.5e-17;
Matches 231; Conservative 0; Mismatches 186; Indels 12; Gaps 1;

```
Qy    366 attgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaataccca 425
          | ||  |    || || || |||||| |     |  | ||    ||  |||| ||| | ||
Db     46 AGTGGCCAGTGGGTCAGGATTCATCGTATCGGAGGATGGACTGATTGTGACAAATGCTCA 105

Qy    426 cgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgc 485
          ||| ||  ||   |    | |  |||| ||   ||| ||  |          |||   ||
Db    106 CGTGGTCACCAACAAAAACCGGGTCAAGGTTGAGCTGAAGAATGGAGCTACCTATGAAGC 165

Qy    486 caaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacgga 545
          |||| |||  | |  |||||   |  | ||  | || ||| | || || |||     ||
Db    166 CAAAATCAAGGATGTGGATGAAAAGGCGGACATTGCGCTTATCAAGATTGACCACCAGGG 225

Qy    546 agagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgc 605
          | |||| || ||| |     |||||    ||  | |  ||| ||| || |||| || | |
Db    226 AAAGCTGCCAGTCCTGCTGCTCGGCCGCTCCTCAGAGCTGAGACCTGGAGAATTTGTAGT 285

Qy    606 tgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaa 665
          |||||| ||    |||||     || | ||||  || ||| | || |||||   | |||
Db    286 TGCCATTGGAAGCCCCTTTTCTCTTCAAAACACAGTCACCACTGGGATCGTCAGCACCAC 345

Qy    666 aggcagaagc------------ctgcccaacgaaagctacacacccttcatccaaaccga 713
            || ||               |      |  | | |    || ||| || || ||
Db    346 CCAGCGAGGCGGCAAAGAGCTGGGACTTCGGAACTCCGATATGGACTACATTCAGACAGA 405

Qy    714 cgttgccatcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgt 773
          || |  |||||||   ||  |||||| |||||| |  | ||| |  | ||  ||||  |
Db    406 CGCTATCATCAATTATGGAGATTCCGGAGGCCCGTTAGTAAACCTGGATGGCGAGGTGAT 465

Qy    774 cggcatcaa 782
            || || ||
Db    466 TGGGATTAA 474
```

RESULT 4
AW742283
LOCUS AW742283 660 bp mRNA EST 07-SEP-2000
DEFINITION up55a10.y1 Soares_mouse_NMIE Mus musculus cDNA clone IMAGE:2779962 5' similar to TR:Q92743 Q92743 NOVEL SERINE PROTEASE. ;, mRNA sequence.
ACCESSION AW742283
VERSION AW742283.1 GI:7654068
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 660)

```
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 431.
FEATURES             Location/Qualifiers
     source          1. .660
                     /organism="Mus musculus"
                     /strain="C3H x 101 (F1 stock)"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:2779962"
                     /clone_lib="Soares_mouse_NMIE"
                     /sex="male"
                     /dev_stage="newborn"
                     /lab_host="DH10B"
                     /note="Organ: inner ear, 170 pooled; Vector: pT7T3D-Pac;
                     Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed
                     with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCACACTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not
                     I and Eco RI sites of the modified pT7T3 vector. Library
                     is normalized, and was constructed and donated by Bento
                     Soares and M.Fatima Bonaldo (University of Iowa) and R.
                     Hardisty, A. Varela-Carver, P. Mburu and S.D.M. Brown (MRC
                     UK Mouse Genome Centre and Mammalian Genetics Unit,
                     Harwell, UK)."
BASE COUNT      188 a    154 c    176 g    141 t      1 others
ORIGIN
```

```
  Query Match             7.0%;  Score 97.4;  DB 94;  Length 660;
  Best Local Similarity   53.8%;  Pred. No. 5.9e-17;
  Matches  231;  Conservative    0;  Mismatches  186;  Indels   12;  Gaps    1;

Qy    366 attgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaataccca 425
          | ||  |   || || || |||||| |     |  | ||    ||  |||| ||| | ||
Db     60 AGTGGCCAGTGGGTCAGGATTCATCGTATCGGAGGATGGACTGATTGTGACAAATGCTCA 119

Qy    426 cgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgc 485
          ||| ||  ||   |    | |  |||| ||   ||| ||  |           |||   ||
Db    120 CGTGGTCACCAACAAAAACCGGGTCAAGGTTGAGCTGAAGAATGGAGCTACCTATGAAGC 179

Qy    486 caaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacgga 545
          |||| |||  | |  |||||   |  | ||  | || ||| | || || |||     ||
Db    180 CAAAATCAAGGATGTGGATGAAAAGGCGGACATTGCGCTTATCAAGATTGACCACCAGGG 239

Qy    546 agagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgc 605
          | |||| || ||| |     |||||    ||  | |  ||| ||| || ||||  || |
Db    240 AAAGCTGCCAGTCCTGCTGCTCGGCCGCTCCTCAGAGCTGAGACCTGGAGAATTTGTAGT 299

Qy    606 tgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaa 665
          |||||| ||    |||||     || | ||||  || ||| | || ||||||   | |||
Db    300 TGCCATTGGAAGCCCCTTTTCTCTTCAAAACACAGTCACCACTGGGATCGTCAGCACCAC 359

Qy    666 aggcagaagc----------ctgcccaacgaaagctacacacccttcatccaaaccga 713
              || ||               |      |   | | |    || ||| || || ||
Db    360 CCAGCGAGGCGGCAAAGAGCTGGGACTTCGGAACTCCGATATGGACTACATTCAGACAGA 419

Qy    714 cgttgccatcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgt 773
          || |  |||||||   || |||||| | |||||| |  | ||| |  | ||  |||| |
Db    420 CGCTATCATCAATTATGGAAATTCCTGAGGCCCGTTAGTAAACCTGGATGGCGAGGTGAT 479

Qy    774 cggcatcaa 782
          || || ||
Db    480 TGGGATTAA 488
```

```
RESULT   5
AI159029
LOCUS       AI159029      456 bp    mRNA            EST       02-OCT-1998
DEFINITION  vz83a08.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
            IMAGE:1333046 5' similar to TR:Q92743 Q92743 NOVEL SERINE PROTEASE.
            ;, mRNA sequence.
ACCESSION   AI159029
VERSION     AI159029.1  GI:3692211
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:692590
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 241.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1333046"
                     /clone_lib="Soares_mammary_gland_NbMMG"
                     /sex="male"
                     /tissue_type="mammary gland"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT      124 a    107 c    127 g     97 t      1 others
ORIGIN
```

```
  Query Match             6.8%;  Score 95.4;  DB 16;  Length 456;
  Best Local Similarity   53.9%;  Pred. No. 2e-16;
  Matches  226;  Conservative    0;  Mismatches  181;  Indels   12;  Gaps    1;

Qy    376 ggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgcc 435
          || || || |||||| |     |  | ||    ||  |||| ||| | ||||| ||  ||
Db     10 GGGTCAGGATTCATCGTATCGGAGGATGGACTGATTGTGACAAATGCTCACGTGGTCACC 69

Qy    436 ggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatc 495
          |    | |  |||| ||   ||| ||  |           |||   |||||| |||
Db     70 AACAAAAACCGGGTCAAGGTTGAGCTGAAGAATGGAGCTACCTATGAAGCCAAAATCAAG 129
```

```
Qy    496 ggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctaccc 555
          | |  |||||   |  | ||  | || ||| | || || |||     || | |||| ||
Db    130 GATGTGGATGAAAAGGCGGACATTGCGCTTATCAAGATTGACCACCAGGGAAAGCTGCCA 189

Qy    556 gtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggc 615
          ||| |     |||||    ||  | |  ||| ||| || |||| || | |||||| ||
Db    190 GTCCTGCTGCTCGGCCGCTCCTCAGAGCTGAGACCTGGAGAATTTGTAGTTGCCATTGGA 249

Qy    616 gcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaaaggcagaagc 675
            |||||     || | ||||  || ||| | || |||||   | |||      || ||
Db    250 AGCCCCTTTTCTCTTCAAAACACAGTCACCACTGGGATCGTCAGCACCACCCAGCGAGGC 309

Qy    676 ------------ctgcccaacgaaagctacacacccttcatccaaaccgacgttgccatc 723
                       |      |   | | |    || ||| || || |||| |  ||||
Db    310 GGCAAAGAGCTGGGACTTCGGAACTCCGATATGGACTACATTCAGACAGACGCTATCATC 369

Qy    724 aatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgtcggcatcaa 782
          |||   || |||||||| |  ||| |  | ||| |  | ||  ||||  | || || ||
Db    370 AATTATGGAAATTCCGGAGCGCCGTTAGTAAACCTGGATGGCGAGGTGATTGGGATTAA 428
```

```
RESULT   6
AW654015
LOCUS       AW654015      521 bp    mRNA            EST       14-JUL-2000
DEFINITION  103226 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   AW654015
VERSION     AW654015.1  GI:7419841
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 521)
  AUTHORS   Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
            Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
            ,W.W. and Keele,J.W.
  TITLE     Design and use of four pooled tissue normalized cDNA libraries for
            EST discovery in cattle
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 103  row: D  column: 5
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .521
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 1BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from lymph node, ovary,
                     fat, hypothalamus, and pituitary."
BASE COUNT      134 a    135 c    147 g    105 t
ORIGIN
```

```
  Query Match             6.8%;  Score 94.6;  DB 93;  Length 521;
  Best Local Similarity   53.1%;  Pred. No. 3.5e-16;
  Matches  233;  Conservative    0;  Mismatches  194;  Indels   12;  Gaps    1;

Qy    366 attgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaataccca 425
          | ||  |   || || || |||||| |   | || | ||    ||  |||| ||  ||||
Db     68 AGTGGCCAGTGGGTCTGGGTTCATCGTGTCCGAAGATGGACTGATTGTGACAAACGCCCA 127

Qy    426 cgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgc 485
          ||| ||  ||   | |    | |||||||  ||| ||  ||      |   |||   ||
Db    128 CGTGGTGACCAACAAGCATCGGGTCAAAGTCGAGCTTAAGAACGGTGCCACCTATGAGGC 187

Qy    486 caaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacgga 545
          |||| |||  |    ||| |   |  | ||  | |||||| | |||||||||      ||
Db    188 CAAAATCAAAGACGTGGACGAGAAGGCGGACATTGCCCTTATCAAAATCGACCACCAGGG 247

Qy    546 agagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgc 605
          | |||| || ||  |      | |||    ||    |  ||   ||||| || |  || |
Db    248 AAAGCTGCCTGTGCTACTGCTTGGCCGCTCCTCGGAGCTGCGGCCGGGAGAGTTCGTGGT 307

Qy    606 tgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaa 665
          |||||  ||    || ||   | || | ||||  || ||| | || |||||   | |||
Db    308 CGCCATTGGAAGCCCGTTTTCCCTTCAAAACACGGTCACCACGGGGATCGTCAGCACCAC 367

Qy    666 aggcag------------aagcctgcccaacgaaagctacacacccttcatccaaaccga 713
                |             |     || |    |     |||    || |||||| || ||
Db    368 CCAGCGTGGTGGCAAAGAACTGGGGCTCCGGAACTCGGACATGGACTACATCCAGACAGA 427

Qy    714 cgttgccatcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcgt 773
          ||   ||||||    || || ||||| |||||   |  | ||| |  | ||  | || |
Db    428 CGCCATCATCAACTATGGAAACTCCGGAGGCCCATTAGTAAACCTGGATGGTGAAGTGAT 487

Qy    774 cggcatcaattcgcaaata 792
          || || ||  | |  | |
Db    488 TGGAATTAACACACTGAAA 506
```

```
RESULT   7
W17262
LOCUS       W17262        474 bp    mRNA            EST       29-APR-1996
DEFINITION  zb17a08.r1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
            IMAGE:302294 5' similar to SW:HTRA_SALTY P26982 PROTEASE DO
            PRECURSOR ;, mRNA sequence.
ACCESSION   W17262
VERSION     W17262.1  GI:1291697
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 474)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: ETPrimer
            High quality sequence stop: 446.
FEATURES             Location/Qualifiers
     source          1. .474
                     /organism="Homo sapiens"
                     /db_xref="GDB:1247226"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:302294"
                     /clone_lib="Soares_fetal_lung_NbHL19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
```

strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbHH19W."

BASE COUNT 122 a 121 c 135 g 93 t 3 others
ORIGIN

Query Match 6.6%; Score 92.2; DB 146; Length 474;
Best Local Similarity 57.4%; Pred. No. 1.6e-15;
Matches 166; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

```
Qy    376 ggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgcc 435
          || || || || ||  |     || | ||    ||| |||| ||| ||||||| ||  ||
Db     41 GGGTCTGGGTTTATTGTGTCGGAAGATGGACTGATCGTGACAAATGCCCACGTGGTGACC 100

Qy    436 ggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatc 495
            |  | |  |||||||   ||| ||  ||     |   ||    |||||| |||
Db    101 AACAAGCACCGGGTCAAAGTTGAGCTGAAGAACGGTGCCACTTACGAAGCCAAAATCAAG 160

Qy    496 ggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctaccc 555
          | |  |||||   || | ||  |||| ||  | ||||| |||     ||   |||| ||
Db    161 GATGTGGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCACCAGGGCAAGCTGCCT 220

Qy    556 gtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggc 615
          |||  |     | |||    ||  | |  ||   ||||| || |  || |  ||||||||
Db    221 GTCCTGCTGCTTGGCCGCTCCTCAGAGCTGCGGCCGGGAGAGTTCGTGGTCGCCATCGGA 280

Qy    616 gcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgcca 664
            || ||   | || | ||||  || ||| |||| ||||||  | |||
Db    281 AGCCCGTTTTCCCTTCAAAACACAGTCACCACCGGGATCGTGAGCACCA 329
```

RESULT 8
R52327
LOCUS R52327 485 bp mRNA EST 18-MAY-1995
DEFINITION yg75a09.r1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:38967 5' similar to SP:HTRA_SALTY P26982 PROTEASE DO PRECURSOR ;, mRNA sequence.
ACCESSION R52327
VERSION R52327.1 GI:814229
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 485)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1476
High quality sequence stops: 391 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1476 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 391.
FEATURES Location/Qualifiers
source 1. .485
/organism="Homo sapiens"
/db_xref="GDB:411508"
/db_xref="taxon:9606"
/clone="IMAGE:38967"
/clone_lib="Soares infant brain 1NIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not I; Site_2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 126 a 121 c 135 g 103 t
ORIGIN

Query Match 6.6%; Score 92.2; DB 144; Length 485;
Best Local Similarity 57.4%; Pred. No. 1.6e-15;
Matches 166; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

```
Qy    376 ggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgcc 435
          || || || || ||  |     || | ||    ||| |||| ||| ||||||| ||  ||
Db      5 GGGTCTGGGTTTATTGTGTCGGAAGATGGACTGATCGTGACAAATGCCCACGTGGTGACC 64

Qy    436 ggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatc 495
            |  | |  |||||||   ||| ||  ||     |   ||    |||||| |||
Db     65 AACAAGCACCGGGTCAAAGTTGAGCTGAAGAACGGTGCCACTTACGAAGCCAAAATCAAG 124

Qy    496 ggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctaccc 555
          | |  |||||   || | ||  |||| ||  | ||||| |||     ||   |||| ||
Db    125 GATGTGGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCACCAGGGCAAGCTGCCT 184

Qy    556 gtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggc 615
          |||  |     | |||    ||  | |  ||   ||||| || |  || |  ||||||||
Db    185 GTCCTGCTGCTTGGCCGCTCCTCAGAGCTGCGGCCGGGAGAGTTCGTGGTCGCCATCGGA 244

Qy    616 gcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgcca 664
            || ||   | || | ||||  || ||| |||| ||||||  | |||
Db    245 AGCCCGTTTTCCCTTCAAAACACAGTCACCACCGGGATCGTGAGCACCA 293
```

RESULT 9
AW660844
LOCUS AW660844 504 bp mRNA EST 14-JUL-2000
DEFINITION 99933 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AW660844
VERSION AW660844.1 GI:7426671
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 504)
AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid,W.W. and Keele,J.W.
TITLE Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred

```
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 101  row: N  column: 22
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES              Location/Qualifiers
     source           1. .504
                      /organism="Bos taurus"
                      /db_xref="taxon:9913"
                      /clone_lib="MARC 1BOV"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                      Library made from pooled tissue from lymph node, ovary,
                      fat, hypothalamus, and pituitary."
BASE COUNT      113 a    143 c    146 g    102 t
ORIGIN


  Query Match             6.5%;  Score 91;  DB 93;  Length 504;
  Best Local Similarity   56.5%;  Pred. No. 3.6e-15;
  Matches  169;  Conservative    0;  Mismatches  130;  Indels    0;  Gaps    0;

Qy    366 attgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaataccca 425
          | ||  |   || || || |||||| |   | || | ||    ||  |||| ||  ||||
Db    195 AGTGGCCAGTGGGTCTGGGTTCATCGTGTCCGAAGATGGACTGATTGTGACAAACGCCCA 254

Qy    426 cgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgc 485
          ||| ||  ||   | |    | ||||||||  ||| ||  ||     |   |||   ||
Db    255 CGTGGTGACCAACAAGCATCGGGTCAAAGTCGAGCTTAAGAACGGTGCCACCTATGAGGC 314

Qy    486 caaactcatcggttcggatgtccaatccgatgtcgccccttctgaaaatcgacgcaacgga 545
          |||| |||  |    |||||   | || | |||||| | |||||||||     ||
Db    315 CAAAATCAAAGACGTGGATGAGAAGGCGGACATTGCCCTTATCAAAATCGACCACCAGGG 374

Qy    546 agagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgc 605
          | |||| |||||  |    | |||    ||    |  ||   ||||| || |  || |
Db    375 AAAGCTGCCCGTGCTACTGCTTGGCCGCTCCTCGGAGCTGCGGCCGGGAGAGTTCGTGGT 434

Qy    606 tgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgcca 664
           ||||| ||    || ||   | || | ||||  || ||| | || |||||   | |||
Db    435 CGCCATTGGAAGCCCGTTTTCCCTTCAAAACACGGTCACCACGGGGATCGTCAGCACCA 493


RESULT  10
F07134
LOCUS       F07134        350 bp    mRNA            EST       20-FEB-1995
DEFINITION  HSC1WG081 normalized infant brain cDNA Homo sapiens cDNA clone
            c-1wg08, mRNA sequence.
ACCESSION   F07134
VERSION     F07134.1  GI:672779
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 350)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
            ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
            Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: molecular integration of the analysis of the human genome
            and its expression
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
  MEDLINE   95277534
COMMENT     Contact: Genethon
            Genexpress-Genethon
            Genethon Centre de recherche sur le Genome Humain
            1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
            Tel: 33169472800
            Fax: 33160778698
            Email: genexpress@genethon.fr
            Single read.
            Genexpress_library_idt: C; Genexpress_sequence_idt: ylc-1wg08
            Seq primer: (-21)M13_universal.
FEATURES              Location/Qualifiers
     source           1. .350
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="c-1wg08"
                      /clone_lib="normalized infant brain cDNA"
                      /sex="Female"
                      /tissue_type="total brain"
                      /dev_stage="3 months old"
                      /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
                      Site_2: NotI; sex=Female; dev_stage=3 months old;
                      isolate=muscular atrophy patient;  tissue_type=total brain
                      ;  total mRNA was oligo-(dT) primed and directionally
                      cloned 5' -> 3' into the  HindIII -> NotI sites of the
                      lafmid BA vector. Clone library from B.Soares,  Psychiatry
                      Dept. Columbia University, USA.  Normalization_method:
                      Bento Soares, P.N.A.S in press"
BASE COUNT       88 a     95 c    100 g     66 t      1 others
ORIGIN


  Query Match             6.4%;  Score 89.6;  DB 140;  Length 350;
  Best Local Similarity   56.7%;  Pred. No. 8.2e-15;
  Matches  164;  Conservative    0;  Mismatches  125;  Indels    0;  Gaps    0;

Qy    376 ggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgcc 435
          || || || || ||  |     || | ||    ||| |||| | | || |||| ||  ||
Db      6 GGGTCTGGGTTTATTGTGTCGGAAGATGGTCTGATCGTGACAATTGCCNACGTGGTGACC 65

Qy    436 ggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatc 495
          | |  | |  |||||||   ||| ||  ||      |   ||    |||||| |||
Db     66 AACAAGCACCGGGTCAAAGTTGAGCTGAAGAACGGTGCCACTTACGAAGCCAAAATCAAG 125

Qy    496 ggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctaccc 555
          | |  |||||   || | ||  |||| ||  | ||||| |||     ||   |||| ||
Db    126 GATGTGGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCACCAGGGCAAGCTGCCT 185

Qy    556 gtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggc 615
          ||| |     | |||    ||  | |  ||   ||||| || |  || |  ||||||||
Db    186 GTCCTGCTGCTTGGCCGCTCCTCAGAGCTGCGGCCGGGAGAGTTCGTGGTCGCCATCGGA 245

Qy    616 gcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgcca 664
            || ||   | || | ||||  || ||| |||| ||||||   | |||
Db    246 AGCCCGTTTTCCCTTCAAAACACAGTCACCACCGGGATCGTGAGCACCA 294


RESULT  11
AV614726
LOCUS       AV614726      529 bp    mRNA            EST       30-AUG-2000
DEFINITION  AV614726 Bos taurus adipocyte cell line Bos taurus cDNA clone
            E1AD006H06 5', mRNA sequence.
ACCESSION   AV614726
VERSION     AV614726.1  GI:9750396
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
            Suzuki,H.
  TITLE     bovine cDNA sequencing
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
```

```
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES              Location/Qualifiers
     source           1. .529
                      /organism="Bos taurus"
                      /db_xref="taxon:9913"
                      /clone="E1AD006H06"
                      /clone_lib="Bos taurus adipocyte cell line"
                      /cell_type="an adipocyte cell line"
                      /lab_host="DH10B"
                      /note="Vector: pZL1; Site_1: Sal1; Site_2: Not1; Poly A
                      was deleted from a Not1 site"
BASE COUNT      132 a    139 c    149 g    108 t      1 others
ORIGIN


  Query Match             6.4%;  Score 89;  DB 37;  Length 529;
  Best Local Similarity   53.0%;  Pred. No. 1.4e-14;
  Matches  222;  Conservative    0;  Mismatches  185;  Indels   12;  Gaps    1;

Qy     366 attgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaataccca 425
           | ||  |   || || || |||||| |   | || | ||    ||  |||| ||  ||||
Db     110 AGTGGCCAGTGGGTCTGGGTTCATCGTGTCCGAAGATGGACTGATTGTGACAAACGCCCA 169

Qy     426 cgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgc 485
           ||| ||  ||   | |     |  |||||||  ||| ||  |       |   |||   ||
Db     170 CGTGGTGACCAACAAGCATCGGGTCAAAGTCGAGCTTAAGAATGGTGCCACCTATGAGGC 229

Qy     486 caaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacgga 545
           |||| |||  |    ||| |   |  | ||  | |||||| | |||||||||     ||
Db     230 CAAAATCAAAGACGTGGACGAGAAGGCGGACATTGCCCTTATCAAAATCGACCACCAGGG 289

Qy     546 agagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgc 605
           | |||| || ||  |     | |||    ||    |  ||   ||||| || |  || |
Db     290 AAAGCTGCCTGTGCTACTGCTTGGCCGCTCCTCGGAGCTGCGGCCGGGAGAGTTCGTGGT 349

Qy     606 tgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgccaa 665
           |||||  ||    || ||   | || | ||||  || ||| | || |||||   | |||
Db     350 CGCCATTGGAAGCCCGTTTTCCCTTCAAAACACGGTCACCACGGGGATCGTCAGCACCAC 409

Qy     666 aggcag----------aagcctgcccaacgaaagctacacacccttcatccaaaccga 713
               |          |     || |    |     |||    || ||||| || ||
Db     410 CCAGCGTGGTGGCAAAGAACTGGGGCTCCGGAACTCGGACATGGACTACATCCAGACAGA 469

Qy     714 cgttgccatcaatccgggcaattccggcggcccgctgttcaacttaaaaggacaggtcg 772
           ||    ||||||    || || ||||| |||||  |  | ||| |  | ||  | || |
Db     470 CGCCATCATCAACTATGGAAACTCCGGAGGCCCATTAGTAAACCTGGATGGTGAAGTGG 528


RESULT  12
AV599976
LOCUS       AV599976      559 bp    mRNA            EST       29-AUG-2000
DEFINITION  AV599976 Bos taurus cartilage fetus Bos taurus cDNA clone
            E1CA047C06 5', mRNA sequence.
ACCESSION   AV599976
VERSION     AV599976.1  GI:9718721
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
            Suzuki,H.
  TITLE     bovine cDNA sequencing
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES              Location/Qualifiers
     source           1. .559
                      /organism="Bos taurus"
                      /db_xref="taxon:9913"
                      /clone="E1CA047C06"
                      /clone_lib="Bos taurus cartilage fetus"
                      /tissue_type="cartilage"
                      /dev_stage="fetus"
                      /lab_host="DH10B"
                      /note="Vector: pZL1; Site_1: Sal1; Site_2: Not1; Poly A
                      was deleted from a Not1 site"
BASE COUNT      135 a    158 c    158 g    108 t
ORIGIN


  Query Match             6.3%;  Score 87.8;  DB 37;  Length 559;
  Best Local Similarity   55.9%;  Pred. No. 3.1e-14;
  Matches  167;  Conservative    0;  Mismatches  132;  Indels    0;  Gaps    0;

Qy     366 attgaacttcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaataccca 425
           | ||  |   || || || |||||| |   | || | ||    ||  |||| ||  ||||
Db     168 AGTGGCCAGTGGGTCTGGGTTCATCGTGTCCGAAGATGGACTGATTGTGACAAACGCCCA 227

Qy     426 cgtcgttgccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgc 485
           ||| ||  ||   | |     |  |||||||  ||| ||  ||      |   |||   ||
Db     228 CGTGGTGACCAACAAGCATCGGGTCAAAGTCGAGCTTAAGAACGGTGCCACCTATGAGGC 287

Qy     486 caaactcatcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacgga 545
           |||| |||  |    ||| |   |  | ||  | |||||| | |||||||||     ||
Db     288 CAAAATCAAAGACGTGGACGAGAAGGCGGACATTGCCCTTATCAAAATCGACCACCAGGG 347

Qy     546 agagctacccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgc 605
           | |||| || ||  |     | |||    ||    |  ||   ||||| || |  || |
Db     348 AAAGCTGCCTGTGCTACTGCTTGGCCGCTCCTCGGAGCTGCGGCCGGGAGAGTTCGTGGT 407

Qy     606 tgccatcggcgcgcccttcggctttgacaacagcgtgaccgccggcatcgtgtccgcca 664
           |||||  ||    || ||   | || | ||||  || ||| | || |||||   | |||
Db     408 CGCCATTGGAAGCCCGTTTTCCCTTCAAAACACGGTCACCACGGGGATCGTCAGCACCA 466


RESULT  13
AI903651/c
LOCUS       AI903651      337 bp    mRNA            EST       30-MAR-2000
DEFINITION  QV-BT032-190299-136 BT032 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI903651
VERSION     AI903651.1  GI:6494038
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 337)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
```

```
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=QV&t2=QV-BT032-136.html
            &t3=190299&t4=1)
            Seq primer: puc 18 forward.
FEATURES              Location/Qualifiers
     source           1. .337
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="BT032"
                      /sex="female"
                      /dev_stage="Adult"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No. 196
                      ,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
BASE COUNT       61 a    118 c     99 g     59 t
ORIGIN


  Query Match             6.2%;  Score 86.2;  DB 26;  Length 337;
  Best Local Similarity   55.9%;  Pred. No. 7.6e-14;
  Matches  185;  Conservative    0;  Mismatches  143;  Indels    3;  Gaps    1;

Qy     339 cccccaagaagaagcagatgacggcggattgaacttcggttcgggcttcatcatcagcaa 398
           ||| || |  |  |   |   ||  |    |  | | || ||||||||     |   ||
Db     330 CCCACACGGTGTCGTCCAGCGCGAGGCGCAGTCCCTGGGCTCGGGCTTATCATTTGCCAG 271

Qy     399 aaacggctacatcctgaccaatacccacgtcgttgccggtatgggcagtatcaaagtcct 458
           | |||||| | |||||||| | ||   | || || |     | |   ||||  |||
Db     270 TGAAGGCTACGTGCTGACCAACAACCCGATGGTCGCGGACGCCGACGAGATCATCGTCTG 211

Qy     459 gctcaacgacaagcgcgaatataccgccaaactcatcggttcggatgtccaatccgatgt 518
           ||    | |    ||||       |||||| |  ||||  | ||    |   |||| ||
Db     210 CCTGTCGGTCCGCAGCGAGCTGCAGGCCAAATTGGTCGGCACCGACCCGCGTACCGACGT 151

Qy     519 cgcccttctgaaaatcgacgcaacggaagagctacccgtcgtcaaaatcggcaatcccaa 578
           |||||  |||||  ||||  |  |   |  | || ||  |||| ||  |  ||  ||  | |
Db     150 GGCCCTGCTGAAGGTCGAGGGCA---AGAACCTGCCGATCGTGAAGCTGGGTGATTGCGA 94

Qy     579 aaatttgaaaccgggcgaatgggtcgctgccatcggcgcgcccttcggctttgacaacag 638
           ||  | ||   ||| || |||||    ||||||||  |||| ||||||||| ||| |
Db      93 GAAGCTCAAGGTGGGTGAGTGGGTACTGGCCATCGGTTCGCCGTTCGGCTTCGACCATTC 34

Qy     639 cgtgaccgccggcatcgtgtccgccaaaggc 669
           |||   |  || |||||   |||||| |||
Db      33 GGTGGGGGGGGGTATCGTCAGCGCCAAGGGC 3


RESULT  14
R22703
LOCUS       R22703         518 bp    mRNA            EST       19-APR-1995
DEFINITION  yg08e06.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:31324 5' similar to SP:HTRA_SALTY P26982 PROTEASE DO
            PRECURSOR ;, mRNA sequence.
ACCESSION   R22703
VERSION     R22703.1  GI:777460
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1522
            High quality sequence stops: 453 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1522   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 453.
FEATURES              Location/Qualifiers
     source           1. .518
                      /organism="Homo sapiens"
                      /db_xref="GDB:403671"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:31324"
                      /clone_lib="Soares infant brain 1NIB"
                      /sex="female"
                      /dev_stage="73 days post natal"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                      I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                      I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3'];
                      double-stranded cDNA was ligated to Hind III adaptors
                      (Pharmacia), digested with Not I and directionally cloned
                      into the Not I and Hind III sites of the Lafmid BA vector.
                      Library went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      131 a    124 c    147 g    112 t      4 others
ORIGIN


  Query Match             6.0%;  Score 84.2;  DB 144;  Length 518;
  Best Local Similarity   56.4%;  Pred. No. 3.2e-13;
  Matches  155;  Conservative    0;  Mismatches  120;  Indels    0;  Gaps    0;

Qy     376 ggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttgcc 435
           || || || || ||  |     || | ||    ||| |||| ||| ||||||| ||  ||
Db       2 GGGTCTGGGTTTATTGTGTCGGAAGATGGACTGATCGTGACAAATGCCCACGTGGTGACC 61

Qy     436 ggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactcatc 495
           | |  | |  |||||||   ||| ||  ||     |   ||    ||||||  |||
Db      62 AACAAGCACCGGGTCAAAGTTGAGCTGAAGAACGGTGCCACTTACGAAGCCAAAATCAAG 121

Qy     496 ggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctaccc 555
           | |  |||||   || | ||  |||| ||  | ||| | |||     ||   |||| ||
Db     122 GATGTGGATGAGAAAGCAGACATCGCACTCATCAAANTTGACCACCAGGGCAAGCTGCCT 181

Qy     556 gtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcggc 615
           ||| |     | |||    ||  | |  ||   ||||| || |  || | ||||||||
Db     182 GTCCTGCTGCTTGGCCGCTCCTCAGAGCTGCGGCCGGGAGANTTCGTGGTCGCCATCGGA 241

Qy     616 gcgcccttcggctttgacaacagcgtgaccgccgg 650
           || ||   | || | ||||  || ||| ||||
Db     242 AGCCCGTTTTCCCTTCAAAACACAGTCACCACCGG 276


RESULT  15
T80106
LOCUS       T80106         288 bp    mRNA            EST       08-MAR-1995
DEFINITION  yd04h09.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:24795 5' similar to SP:HTRA_SALTY P26982 PROTEASE DO
            PRECURSOR ;, mRNA sequence.
ACCESSION   T80106
VERSION     T80106.1  GI:698615
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 288)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1530
            High quality sequence stops: 143 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Insert Length: 1530   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 143.
FEATURES             Location/Qualifiers
     source          1. .288
                     /organism="Homo sapiens"
                     /db_xref="GDB:397142"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:24795"
                     /clone_lib="Soares infant brain 1NIB"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       68 a     71 c     75 g     61 t     13 others
ORIGIN


  Query Match             5.8%;  Score 81.2;  DB 146;  Length 288;
  Best Local Similarity   53.7%;  Pred. No. 1.9e-12;
  Matches  152;  Conservative    0;  Mismatches  131;  Indels    0;  Gaps    0;

Qy    374 tcggttcgggcttcatcatcagcaaaaacggctacatcctgaccaatacccacgtcgttg 433
          | |||  ||| ||  |     | || |  | ||      |||| ||| |||||||
Db      5 TGGGTCTGGGTTTATTGTGTCGGAAGATGGACTGATCGGTGACAAATGCCCACGTGGTGA 64

Qy    434 ccggtatgggcagtatcaaagtcctgctcaacgacaagcgcgaatataccgccaaactca 493
          ||   | |  | |  |||||||   ||| ||  ||     |   ||    |||||| |||
Db     65 CCAACAAGCACCGGGTCAAAGTTGAGCTGAAGAACGGTGCCACTTACGAAGCCAAAATCA 124

Qy    494 tcggttcggatgtccaatccgatgtcgcccttctgaaaatcgacgcaacggaagagctac 553
           | |  |||||   || | ||  |||| ||  | ||||| |||     ||   |||| |
Db    125 AGGATGTGGATGAGAAAGCAGACATCGCACTCATCAAAATTGACCACCAGGGCAAGCTNC 184

Qy    554 ccgtcgtcaaaatcggcaatcccaaaaatttgaaaccgggcgaatgggtcgctgccatcg 613
          | ||| |     |  ||    ||  | |  ||    | || || |  || |   ||||||
Db    185 CTGTCCTGCTNCTTTGCCNCTCCTCAGAGCTGCGGNCNGGAGAGTTCGTNGTCNCCATCG 244

Qy    614 gcgcgcccttcggctttgacaacagcgtgaccgccggcatcgt 656
          |    || ||   | || | | ||  || | | ||||  ||||
Db    245 GANGCCCNTTTTCCCTTCAAATCACAGTNANCNCCGGGNTCGT 287


Search completed: March 22, 2001, 03:35:13

Job time: 4478 sec
```